PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 5 :
C07H 21/00, C12N 15/00, 5/00

A1

(11) International Publication Number: WO 94/21661

(43) International Publication Date: 29 September 1994 (29.09.94)

(21) International Application Number: PCT/US94/03125

(22) International Filing Date: 23 March 1994 (23.03.94)

(30) Priority Data:

08/035,457	23 March 1993 (23.03.93)	US
08/205,716	4 March 1994 (04.03.94)	US

(71) Applicant: THE BOARD OF TRUSTEES OF THE LELAND STANFORD JUNIOR UNIVERSITY [US/US]; Suite 350, 900 Welch Road, Stanford, CA 94303 (US).

(72) Inventors: BLAU, Helen; 580 Cotton Street, Menlo Park, CA 94025 (US). RASTINEJAD, Farzan; 1285 Breezy Lane, Atlanta, GA 30329 (US).

(74) Agents: ROWLAND, Bertram, I. et al.; Flehr, Hohbach, Test, Albritton & Herbert, 4 Embarcadero Center, Suite 3400, San Francisco, CA 94111-4187 (US).

(81) Designated States: AU, CA, JP, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

Published
With international search report.

(54) Title: CELLULAR REGULATION WITH RIBOREGULATORS

(57) Abstract

UTRs of cell structural proteins, particularly associated with muscle, or their complementary sequences or ribozymes comprising complementary sequences ("riboregulators") are shown to be capable of regulation of cell division and/or cell differentiation. Thus by providing the sequences exogenously or providing genes which transcribe the sequences endogenously, the fate of cells can be controlled, as to division or differentiation. The sequences can be used for controlling proliferation of neoplastic cells or in gene therapy, or to treat diseases associated with defects in cellular differentiation, or to drive cells to differentiate or to divide.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	GB	United Kingdom	MR	Mauritania
AU	Australia	GE	Georgia	MW	Malawi
BB	Barbados	GN	Guinea	NE	Niger
BE	Belgium	GR	Greece	NL	Netherlands
BF	Burkina Faso	HU	Hungary	NO	Norway
BG	Bulgaria	IE	Ireland	NZ	New Zealand
BJ	Benin	IT	Italy	PL	Poland
BR	Brazil	JP	Japan	PT	Portugal
BY	Belarus	KE	Kenya	RO	Romania
CA	Canada	KG	Kyrgystan	RU	Russian Federation
CF	Central African Republic	KP	Democratic People's Republic of Korea	SD	Sudan
CG	Congo			SE	Sweden
CH	Switzerland	KR	Republic of Korea	SI	Slovenia
CI	Côte d'Ivoire	KZ	Kazakhstan	SK	Slovakia
CM	Cameroon	LI	Liechtenstein	SN	Senegal
CN	China	LK	Sri Lanka	TD	Chad
CS	Czechoslovakia	LU	Luxembourg	TG	Togo
CZ	Czech Republic	LV	Latvia	TJ	Tajikistan
DE	Germany	MC	Monaco	TT	Trinidad and Tobago
DK	Denmark	MD	Republic of Moldova	UA	Ukraine
ES	Spain	MG	Madagascar	US	United States of America
FI	Finland	ML	Mali	UZ	Uzbekistan
FR	France	MN	Mongolia	VN	Viet Nam
GA	Gabon				

CELLULAR REGULATION WITH RIBOREGULATORS

This application is a continuation-in-part of application Serial No. 08/035,457, filed March 23, 1993.

INTRODUCTION

Technical Field

The field of this invention is proliferation and differentiation regulation.

Background

A complex regulatory circuitry underlies mammalian growth and differentiation. Although many tissue-specific regulators are transcription factors that bind DNA, others act indirectly. For example, regulators can sequester other regulators in intracellular compartments, bind cooperatively with ubiquitous transcription factors, modify these factors by changing their phosphorylation state, or facilitate their dimerization. Thus, an elucidation of growth and differentiation control pathways requires a combination of assays based on binding to known molecules with assays based on function.

A genetic complementation approach provides a means of identifying regulators based on function without preconceptions as to their nature. Major obstacles to the application of a genetic complementation approach to mammalian differentiation have included devising a means of identifying appropriate cells to complement and a method for recovering the DNA once complementation is achieved. Moreover, regulators of differentiation may be

inhibitory to growth which is essential to isolation of complemented cells by drug selection.

The regulation of growth, differentiation and viral replication and transformation can have many applications for prophylaxis and therapy. There are many indications where the uncontrolled growth of cells is involved with severe diseases, such as cancer, psoriasis, hyperthyroidism, and the like. Problems in treating these various diseases have been associated with selective regulation of cells which are normally non-dividing, as compared to other cells which are dividing normally and are necessary to the health of the host. Also, there are many situations where failure to differentiate can lead to genetic diseases or the inability of the host to respond to infection. In the case of viral diseases or viral transformation of a cell, the disease is predicated upon substantially uncontrolled proliferation of the virus in a cell and/or modification of the cell genome. There is, therefore, substantial interest in finding opportunities which will allow for control of biological proliferation, particularly cell division and/or differentiation.

Relevant Literature

For reviews of articles associated with the genetic complementation approach to identify regulators, see, for example, Herskowitz (1985), Cold Spring Harbor Symposium, *Quant. Biol.* 50, 565-574; Nusslein-Volhard (1991) *Development Supplement* 1, 1-10; Surana et al. (1991) *Cell* 65, 145-161; Simon et al. (1991) *Cell* 67, 701-716; Blau (1988) *Cell* 53, 673-674; Pan et al. (1992) *Som. Cell & Mol. Genet.* 18, 163-177; Simmons et al. (1992) *J. Immun.* 148, 267-271; Durkop et al. (1992) *Cell* 68, 421-427; Greco et al. (1987) *Proc. Natl. Acad. Sci. USA* 84, 1565-1569; Schatz and Baltimore (1988) *Cell* 53, 107-115; Kitayama et al. (1989) *Cell* 56, 77-84; Velazquez et al. (1992) *Cell* 70, 313-322.

Numerous experiments suggest that the differentiated state requires continuous regulation by both positive and negative regulators: Blau et al. (1983) *Cell* 32, 1171-1180; Blau et al. (1985) *Science* 230, 758-766; Wright (1984) *J. Cell Biol.* 98, 436-443; Baron and Maniatis (1986) *Cell* 46, 591-602; Spear and Tilghman (1990) *Mol. Cell. Biol.* 10, 5047-5054; Wu et al.

(1991) *Mol. Cell. Biol.* 11, 4423-4430; Way and Chalfie (1989) *Genes Dev.* 3, 1823-1833; Belote and Baker (1987) *Proc. Natl. Acad. Sci. USA* 84, 8026-8030; Blau and Baltimore(1991) *J. Cell. Biol.* 112, 781-783; Blau (1992) *Annual Review of Biochemistry* 61, 1213-1230.

In muscle tissue, it is found that growth factors prevent the expression of myogenic genes (Vaidya et al. (1991) *J. Cell. Biol.* 114, 809-820; Li et al. (1992) *Genes Dev.* 6, 676-689); whereas inhibitors of cell growth promote differentiation (Silberstein et al. (1986) *Cell* 46, 1075-1081; Gu et al. (1993) *Genes Dev.* 6, 1783-1798.

Most tumor cells display an undifferentiated or poorly differentiated phenotype, suggesting that the pathways of tumor suppression and differentiation are tightly linked (Sachs (1978) *Nature* 274, 535-539; Sachs (1987) *Cancer Research* 47, 1981-1986; Mintz and Fleischman (1981) *Adv. Cancer Res.* 34, 211-278; Harris (1985) *J. Cell Sci.* 79, 83-94). Suppression of tumor formation by introduction of suppressor genes often correlates with reappearance of the differentiated phenotype (Francis-Lang et al. (1992) *Mol. Cell. Biol.* 12, 5793-5800; Harris (1985) *supra*; Murate et al. (1984) *Proc. Natl. Acad. Sci. USA 81,* 3394-3398; Peehl and Stanbridge (1982) *Int. J. Cancer* 30, 113-120).

Inactivation of tumor suppressor genes and appearance of the tumorigenic phenotype is well correlated with alteration of the cytoskeletal architecture as evidenced by changes in both expression and organization of extracellular matrix and microfilament components (Chan et al. (1989) *Proc. Natl. Acad. Sci. USA* 86, 2747-2751; Cooper et al. (1987) *Cancer Research* 47, 4493-4500. Garrels and Franza (1989) *J. Biol. Chem,* 264, 5299-5312; Rodriguez Fernandez et al. (1992) *J. Cell. Biol.* 119, 427-438; Gluck et al. (1993) *Proc. Natl. Acad. Sci. USA* 90, 383-387; and Eiden et al. (1991 *Mol Cell. Biol.* 11, 5321-5329. Differentiation-associated components of the cytoskeletal apparatus are among the genes that decline dramatically in neoplastic cells and return to normal levels in cells that revert back to the normal phenotype (Leavitt et al. (1986) *Mol. Cell. Biol.* 6, 2721-2726; Cooper et al. (1985) *Mol. Cell. Biol.* 5, 972-983; Lin et al. (1985) *J. Cell.*

Biol. 100, 692-703; Matsumura et al. (1983) *J. Biol. Chem.* 258, 13954-13964; Leavitt et al. (1982) *Cell* 28, 259-268).

Cis-regulation of actin genes by 3'UTRs has been reported, but not trans regulation with the exception of Lloyd and Gunning (1993) *J. Cell Biol.* 121, 73-82 who note a change in morphology upon expression of one *γ*-actin 3'UTR. Expression of a stably transfected chimeric gene of the 5' regulatory region of *β*-actin linked to the 3'UTR of *α*-skeletal actin does not decline during differentiation, but increases like the endogenous *α*-skeletal actin gene (Sharp et al. (1989) *Gene* 80, 293-304). Conversely, when the 3'UTR of *β*-actin is linked to heterologous promoters, expression from these promoters, like the endogenous *β*-actin gene declines during differentiation (DePonti-Zilli et al. (1988) *Proc. Natl. Acad. Sci. USA* 85, 1389-1393). 3'UTR sequences also act in cis to control the location within a cell of actin mRNAs., Kislauskis et al. (1993) *J. Cell Biol.* 123, 165-172. Blau and Rastinejad (1993) *J. Cell Biology Supplement* 17A: 58, report activation of muscle-specific promoter constructs in a mutant cell in trans by the 3' untranslated regions of structural genes. See also, Rastinejad and Blau (1993) *Cell* 72, 903-917 and Rastinejad et al. (1993) *Cell* 75, 1107-1117.

A dsRNA-regulated protein kinase known as PKR (protein kinase activated by RNA) has been described: Clemens et al. (1993) *J. Interferon Res.* ; Petryshyn et al. (1988) *Proc. Natl. Acad. Sci. USA* 13, 241; Chong et al. (1992) *EMBO J.* 11, 1553-1562; Koromilas et al. (1992) *Science* 257, 1685-1689; and Meurs et al. (1993) *Proc. Natl. Acad. Sci. USA* 90, 232-236. Viral and synthetic dsRNA having PKR regulatory activity has been described by Kostura and Mathews (1989) *Mol. Cell Biol.* 9, 1576-1586; Gunnery et al. (1990) *Proc. Natl. Acad. Sci. USA* 87, 8687-8691; Clarke et al. (1991) *Nucl. Acids Res.* 19, 243-248; Roy et al. (1991) *J. Virol.* 65, 632-640; Manche et al. (1992) *Mol. Cell Biol.* 12, 5238-5248.

PKR peak activity in adipocytes has been correlated with the appearance of an uncharacterized poly(A)$^+$ dsRNA (Li and Petryshyn (1991) *Eur. J. Biochem.* 195, 41-48). The secondary structure of the *α*-tropomyosin 3'UTR is similar to the optimal structure for dsRNA PKR activators deduced from studi s of adenovirus or HIV-1 TAR (Mathews and Shenk (1991) *J.*

Virol. 65, 5657-5662; and Roy et al., *ibid* 65, 632-640. No cellular RNA activators of PKR have previously been described.

SUMMARY OF THE INVENTION

The identification and use of transcribed untranslated RNA regions of genes comprising at least one exon (UTRs) is provided, where the UTRs provide cellular regulation. Particularly the UTRs are associated with genes associated with cellular proliferation or differentiation. Genetic complementation is employed for identification of the UTRs associated with regulation of cell division and differentiation, where the UTRs may then be employed for the modulation of cell growth and differentiation, both as to normal and abnormal physiology, as well as viral replication. The sequences may be used to diagnose the neoplastic status of cells. In addition, the sequences can be employed to identify factors to which the sequences bind intracellularly, and as such may be employed in drug design.

The employment of the subject methodology is exemplified with the identification of the 3'UTR of α-tropomyosin, as well as a fragment thereof, as a cellular activator for PKR.

DESCRIPTION OF THE SPECIFIC EMBODIMENTS

Methods and compositions are provided for identification of UTR sequences associated with the regulation of cell division and differentiation, particularly as to regulation of expression of genes associated with cell division and differentiation. The methods employ genetic complementation, where a cell line is produced which is deficient in expression of proteins associated with cell differentiation, so as to lack the ability to differentiate to a more mature cell, while having a relatively high potential for tumorigenic growth as evidenced by anchorage independent cell growth, as well as tumor formation in a viable mammalian host, or other appropriate phenotypes. The cell can serve to screen agents which genetically complement its deficiency.

The cells may then be modified by introduction of DNA constructs providing for transcription of UTRs and then screened for their tumorigenic phenotype.

An alternative method is to employ PKR as a specific binding member for UTRs which have been selected in accordance with a program for defining secondary structure, e.g. free energy minimization according to Zuker, (1989) *Science* 244, 48-52 and Jaeger et al. (1989) *Proc. Natl. Acad. Sci. USA* 86, 7706-7710, viewed with Loopviewer (Gilbert (1990), a Macintosh program for visualizing RNA secondary structure. Published electronically on the Internet, available via anonymous ftp to iubio.bio.indiana.edu.). Various techniques may be employed for the identification of UTRs, for example, using either competitive or non-competitive assays. Having identified a candidate UTR, the UTR may then be prepared by preparing a transcription cassette to produce the UTR in a bacterial host and the lysate or purified UTR used in the assay. Using PKR as the binding agent, a labeled dsRNA may be used to compete with the candidate dsRNA for binding to PKR. Electrophoretic mobility shift assays ("EMSA") may be employed with an antibody to PKR. Where the UTR is labeled, one can identify the ternary complex by observing the binary complex between the UTR and PKR and the ternary complex which includes the anti-PKR. EMSA may also be used in a competitive mode, where the competitive dsRNA is labeled and the effect of the candidate dsRNA on the binding of the competitive dsRNA to the PKR may be determined.

The RNA untranslated regions of the subject invention ("UTRs") are transcribed and are normally associated with at least one exon in the germline DNA. The UTRs are transcribed from genomic DNA and may be processed, either being removed from the transcription product, by splicing, or remaining and forming part of the messenger RNA, as in the case of the 3'UTR. The UTRs may be the 5' untranslated region, an intron, or the 3' untranslated region. Of particular interest are the 3'UTRs of genes associated with specific cellular functions to be described subsequently, more particularly, those sequences forming secondary structure capable of specifically binding to a target protein, e.g. an enzyme. (By "specific binding" is intended an affinity of at least about 10^{-7} M^{-1}.) The UTRs are primarily associated with genes which are translated and may include domains of RNA genes, such as the telomerase primer, RNA associated with

a ribosome, e.g. 5S RNA, tRNA, or the like. The UTRs may be all or a part of the naturally occurring untranslated region associated with an exon. Usually, the naturally occurring UTR will be at least about 100nt, more usually, at least about 150nt. The active portion may be 60 nt or more, frequently at least about 75 nt, and not more than about 500 nt. The UTRs will be associated with sequences that are normally transcribed during cell differentiation or cell division. A minimal sequence may be defined which is a portion of the naturally occurring UTR, which can fulfill at least 75% of the activity of the naturally occurring UTR. The UTRs will form dsRNA, usually a stem and loop structure. The UTRs will normally be associated with regulation of the expression of a number of genes in trans. In light of the function and activity of the UTRs and their complementary sequences, they may be referred to as "riboregulators."

The UTRs found to modulate cell differentiation and/or division may then be used in a variety of ways, particularly as agents for the modulation of cell differentiation and/or division, using the UTR sequence or the complementary sequence. In addition, the UTRs may be used in affinity binding assays for identifying the proteins in addition to PKR associated with the UTRs in their cellular regulation. Finally, the absence of functional UTRs may be determined as evidence of the tumorigenicity or potential for tumorigenicity of cells.

By virtue of the regulatory activity of the UTRs, nucleic acid compositions, proteins and small synthetic molecules which serve as mimetics may find use in this invention. Thus, the range of molecular weights of compounds, may vary from about 0.2 to 5kD for small synthetic or small naturally occurring organic molecules (distinguished from naturally occurring polymeric molecules, such as proteins, nucleic acids and polysaccharides), to naturally occurring polymeric molecules which may range to as high as about 200kD, e.g. IgM. Generally the oligopeptides and oligonucleotides will be at least about 1kD and not greater than about 10kD, while the larger polymeric molecules will usually be at least about 10kD, more usually at least about 30kD.

A. Development of cell line for screening

The first aspect of this invention to be discussed is the development of a cell line for screening for riboregulators. Various mammalian cell types may be used, which are desirably normal cells which can be maintained in culture. For the purposes of the subject description, the method will be exemplified with myogenic cells. The cells may come from any convenient mammalian host tissue which allows for maintenance of the cells in culture. Thus, the cells may be mouse cells, rat cells, primate cells, human cells, rabbit cells, or the like.

The cells are selected for and, therefore, characterized by being defective in cell differentiation. For purposes of convenience, it is desirable that the cells be able to grow in soft agar in an anchorage-independent manner. In this way, these cells can be used to rapidly screen for constructs which induce cell differentiation, for instance, transcription of UTRs providing for cell differentiation. These cells may be defective in the expression of at least one structural protein associated with cellular differentiation, e.g. cellular structure, particularly cytoskeletal structure, such as microfilaments, microtubules, etc.; or other growth and differentiation related proteins, such as adhesion proteins, membrane proteins, growth factors, oncogenes, and the like. These genes include tropomyosin, troponin I, α-, β- and γ-actin, myosin, tubulin, vimentin, vinculin, collagens, desmin, α-actinin, fibronectin, thrombospondin, and the like; or other proteins, such as NF-IL6, IGF-II, myc, mas, and the like. Desirably, these cells may also be characterized by having a high incidence of tumorigenicity when introduced into a mammalian host, generally generating greater than 90% tumors when inoculated with from about 10^5 to 10^7 cells, depending upon the particular cell line. Desirably, the cells should have a low reversion frequency to the untransformed state, usually a reversion frequency of less than about 10^{-7}.

The cells employed for genetic complementation may be available or may be produced from cell lines which are characterized by being susceptible to mutation to produce cells which are readily distinguished from the parent cells for the purposes of the genetic complementation. The parent cells should be immortal, but not transformed, and have a stable phenotype.

Illustrative of characteristics of interest are that the parent cells should not be defective in a differentiation pathway of interest, should be susceptible to mutagenesis to result in formation of neoplastic cells, and should be associated with a low incidence of tumorigenicity when introduced into a mammalian host, as may be evidenced by a low frequency of generating anchorage-independent colonies as shown by growth in soft agar. The cells should have a good growth potential and be readily capable of being maintained in culture for extended periods of time. The cells should be capable of expression of most, if not all, of the cell structural and regulatory proteins characteristic of the particular cell type, for example, in myogenic cells, should be capable of expressing the helix-loop-helix regulators, and should be capable of both cell differentiation and cell division.

As illustrative of such a cell type is the myogenic cell line C2F3. This cell line expresses all four of the myogenic helix-loop-helix regulators and forms myotubes and gives rise to spontaneous non-differentiating variants at a frequency of less than about 10^{-6}. This cell line was subjected to mutagenesis, using a relatively low dose of mutagen that would yield mutants at a detectable frequency, yet relatively few mutations per cell. In addition, the mutations can be point mutations or deletions. Various concentrations of the mutagen or intensities of X-irradiation can be used, so as to screen for inactivation of a well-characterized single gene or other phenotype of interest.

In carrying out the mutagenesis, one can use a low dosage to minimize the number of mutations which occur, using a known gene to titrate the level of mutagenic agent employed. One can follow a phenotypic characteristic to detect whether cell differentiation deficient cells have been produced. Of particular interest as a phenotypic characteristic is growth in culture without cell differentiation. For determining the dosage, with a myogenic cell line as illustrative, a mutagenic dose-response curve can be obtained, both for inactivation of the gene encoding hypoxanthine phosphoribosyl transferase (HPRT) and for the production of anchorage-independent mutants by assaying growth in soft agar. Based on th experience obtained using N-nitrosomethylurea (NMU), a concentration of 12.5 μg/ml of NMU is

effective, which provides a negligible mortality and a mutant frequency of about 10^{-5}. Further, it was found that growth in soft agar is an effective method for selecting for differentiation defective mutants, since in the case of myogenic mutants, the mutants that grew in soft agar failed to form myotubes in liquid tissue culture media. The resulting mutants expressed different combinations of muscle-specific structural genes and regulatory genes of the myogenic helix-loop-helix (bHLH) family.

The NMU Cell Type

The particular cell line which was produced is called NMU2. It is a clonal cell line and does not form myotubes when grown at high density, either in 10% calf serum or 2% horse serum, conditions under which the parent line, C2F3 differentiates into myotubes. The doubling time for NMU2 in 10% calf serum is 12-14 h, similar to that of the parental C2F3. The NMU2 cells are more refractile and less adherent than C2F3 cells. The NMU2 mutant fails to express mRNAs encoding myogenin, myosin heavy chain, α-sarcomeric actins, troponin I and at least one isoform of tropomyosin. NMU2 does express muscle-specific enolase, MyoD, myf5 and cell surface N-CAM proteins. From this spectrum of expression, it may be concluded that a subset of steps in the myogenic regulatory network have been disrupted in the NMU2 mutant. Furthermore, mutagenesis affected a regulatory step and not the structural genes themselves.

Genetic Complementation by Circumventing the Primary Defect

The mutated cells may be used for genetic complementation, conveniently by stably introducing a genetic construct which allows for identification of the presence of a genetic defect by the failure to obtain expression. Expression of the construct results when the genetic defect is circumvented. By "circumvented" is intended that the pathway which has been blocked is unblocked, although the gene which has been inactivated in the pathway, may or may not be complemented. Therefore, circumvented differs from "rescue" in that the gene which has been mutagenized and inactivated may not be complemented and the pathway may not be

completely restored. The pathways of interest for the subject invention are involved with growth or differentiation, where RNA may act as a regulator in trans.

Various genetic constructs will be employed which involve joining DNA sequences of interest. The sequences of interest may be genetic functional sequences such as promoters, enhancers, LTRs, coding sequences, 3' untranslated regions, introns, exons, splicing signals, origins, etc. The sequences may be associated with the formation of the construct, such as restriction sites, primer sites, polylinkers, and the like. Other sequences may be involved in selection, such as antibiotic resistance genes under the transcriptional control of transcriptional initiation and termination regulatory regions functional in the host.

The construct will usually involve an available vector having a prokaryotic origin, a marker gene(s) for selection and one or more convenient restriction sites. The sequence may be inserted by having complementary termini in the vector and the sequence of interest, either blunt ended or overhang ended. These ends may be available or made so by restriction, tailing with terminal deoxytransferase, *in vitro* mutagenesis, primer repair, nucleases, etc. The various sequences may be ligated, expanded by cloning in a prokaryotic host and analyzed using restriction analysis, sequencing, etc. The various steps involving the insertion or deletion of DNA may be followed after each step by analysis or after a number of steps have been carried out.

The differentiation-defective cell may then be used for introduction of a genetic construct, where the construct comprises a promoter of a gene which is no longer expressed in the mutant and associated with cell differentiation and a marker gene, which allows for detection, particularly selection, of cells in which the promoter is activated. Methods of introduction of DNA constructs will be described subsequently. The promoters which are employed will be associated with differentiation or proliferation and may be different with different cell types. For example, promoters which find use in cells, such as the NMU2 cells of the subject invention, where the comparable gene is not expressed, include α-cardiac actin, α-skeletal actin, troponin I, tropomyosin, and myogenin. Various genes

may be employed in conjunction with the promoter. Of particular interest are genes which allow for selection, particularly providing for resistance to an antibiotic. Genes which may be employed include the guanine-xanthine phosphoribosyl transferase, which provides protection from mycophenolic acid, neomycin resistance gene, puromycin resistance gene, hygromycin resistance gene, dihydrofolate reductase gene, and the like. Alternatively, one may use a gene which can provide for selection by physical identification, such as the β-galactosidase gene or a surface membrane protein gene. In the former case, an appropriate substrate can be provided which renders the cell a blue color or generates a fluorescent product in the cell. In the latter case, one can use fluoresent or other labeled anti-surface membrane protein antibodies, where the cells may then be detected and/or isolated by virtue of their physical appearance, e.g. using a fluorescence activated cell sorter.

To further ensure that the test cells are suitable for genetic complementation, the cells may be fused with the parent cells, where the particular gene is expressed. By showing that the marker gene is expressed when the test cells are fused with the parent cell, but not expressed when fused with itself, demonstrates that the construct is functional and that in the presence of the appropriate regulatory factors, can be expressed.

The resulting cell lines may now be used for screening of various constructs for their ability to regulate cell division and/or cell differentiation. Libraries can be employed where the cDNA is inserted downstream from a promoter, where the promoter is functional in the test cell. The library may be from fetal, neonatal, junior ($\leq$12 y) or adult tissue, where the cells are of the desired type, e.g. myogenic, endothelial, fibroblast, epithelial, neuronal, mucosal, cutaneous, hematopoietic, keratinocytes, hepatocytes, adipocytes, chondrocytic, osteogenic, and the like. The test cell can be used for screening any gene which provides for regulation of cell division and/or cell differentiation. For the purposes of the subject invention, there is a particular interest in identifying a 3'UTR as a regulating element. In this instance, one may employ a genomic or cDNA library, where the fragments present in the library are relatively small as compared to th entire germline gene, ven the

cDNA gene, so as to exclude at least about 25%, more usually at least about 75% of the total cDNA other than the 3'UTR. Shortened cDNA sequences can b achieved with oligo dT priming, polymerase chain reaction rescue, and the like. One may use a subtraction library to enrich for regulatory elements associated with the target cell type. Where the gene is known, one can cleave at a site adjacent the terminal codon, use primers in the polymerase chain reaction which recognize the sequence of the terminal codon and poly(A), so as to provide for sequences involved with the 3'UTR, or the like. For introns or other untranslated sequences of a UTR, the germline DNA may be required. Alternatively, in some mRNA transcripts, an intron is retained, which could be identified in the cDNA, once the gene is shown to have regulatory activity, unassociated with the expressed protein. The significant factor is to detect RNA sequences associated with regulation of the growth or differentiation pathway.

The UTR fragments of the gene may then be inserted downstream from a promoter which is functional in the test cell. The promoter may be constitutive or inducible, desirably being a strong promoter, which may be endogenous to the test cell, a viral promoter, or a promoter from a different mammalian cell. Promoters from various viruses may be employed, such as adenovirus, papilloma virus, simian virus 40, and cytomegalovirus. Endogenous and exogenous promoters may include β-actin, β-globin, metallothionein and glucocorticoid inducible promoters, or promoters inducible by relief of repression by tetracycline or the lac repressor. The promoters should be selected so that the transcription factors which are present in the test cell will allow for transcription of the test sequence. Therefore, the promoter should be selected so as to be independent of the transcriptional pathway which involves the test factors.

The construct may also have other functional sequences, such as an origin which allows for replication in prokaryotic and/or eukaryotic hosts. In case of the latter, one may wish to have a relatively high copy number of the construct in the test cell, so that a high level of transcription will be achi ved. Eukaryotic origins will be for the most part viral origins, which may come from a wide variety of viruses, such as simian virus 40,

adenovirus, papilloma virus, Epstein-Barr virus, etc. In addition, one may have one or more markers present, which allow for selection of t st cells in which the subject constructs have been introduced. For the most part, these markers will involve antibiotic resistance, providing for selection different from the selection associated with the regulatory sequence or markers that allow for selection by panning or on the FACS. Thus, markers described previously may also be used.

While subject to greater uncertainty, in some instances the test cells may be yeast, insect or other non-mammalian eukaryotic cells having a genetic deficiency in cell division or cell differentiation. One could use constructs described above, exchanging the mammalian functional promoter for a promoter functional in the non-mammalian test cell. Numerous vectors are available which are functional in yeast and insect cells and provide promoters which are functional in such cells. These cells would provide for a rapid screen, where positive cells, those where the blocked pathway is unblocked, would identify candidates for screening in the mammalian test cells.

Once having developed the construct, it may now be introduced into the test cell in accordance with known ways. Transfection, electroporation, fusion, polybrene, lipofection, calcium phosphate precipitated DNA, retroviral infection or other conventional technique may be employed. The resulting cells may then be expanded and selected for those cells which have the construct. These cells may then be further expanded, subject to the poor cell growth for cells directed into differentiation, and selected for those cells in which the regulatory sequence has been activated, so that the marker associated with the regulation of differentiation and/or cell division is being expressed. Alternatively, one may transfect with a retrovirus carrying the UTR, whereby when the integrated retrovirus is transcribed, the UTR will be produced. This may lead to higher frequency of introduction into the test cells and single copies of the UTR. These cells may be screened by any convenient assay, which provides for detection of the modulation of cell growth or differentiation.

One assay which may be used is a methylcellulose assay with bromodeoxyuridine (BrdU), where BrdU is cytotoxic to growing cells. Those cells which survive would indicate that the construct is directing the cells to differentiate. Cells which survive may be readily recovered from the methylcellulose and grown in liquid medium. Another assay, particularly as a secondary screening assay, which has shown to be effective and useful is the use of growth in soft agar (0.3 to 0.4% thickening agent). Thus, those cells which are anchorage independent and capable of growing in soft agar may then be further screened. Further confirmation may include testing for tumorigenicity, where the cells may be introduced into a susceptible host, e.g. immunosuppressed, and the formation of tumors determined. This test involves injecting cells containing constructs in a sufficient number of cells equivalent to the development of tumors in at least about 90% of the animals with cells where origianlly the native regulatory sequence is not present. Thus, one can use groups of five animals or more, where one group is injected with test cells which are capable of producing tumors in at least 90% of the animals and the other group is injected with the test cells in which a putative regulatory sequence is believed to be present. A reduction in the number of animals in which tumors occur is indicative of the presence of a regulatory sequence which modulates cell division.

Other tests may include a second reporter gene using a different promoter and selectable markeer, expression of a previously silent or inactive gene encoding an endogenous cell surface protein, assays of colony size by visual inspection, assay of cell number by MTT (Mossman (1983) J. of Immun. Methods 65, 55-63) assay of clones, immunofluorescence for detection of a differentiation specific gene product not normally expressed by the mutant cell type, loss of expression of a cell suface protein normally expressed by the mutant cell type, or the like.

The assays which are carried out in a nutrient medium may be varied as to the amount of serum present. The effect of the riboregulators may vary when growing cells in a nutrient medium that has at least about 10% serum, e.g. fetal calf serum, horse serum, human serum, etc. or less than about 1% serum. Therefore, when screening cells or carrying ut tests in

culture, the medium should be varied to identify the response of the riboregulon to a rich or serum deprived medium.

B. Riboregulators

The UTRs of this invention include sequences from members of gene families transcribed during cell differentiation or cell division, particularly associated with cell morphology, adhesion, cell division or cell differentiation. These include genes associated with the cytoskeleton and the cell surface, e.g. microfilaments and microtubules, cell adhesion proteins, contractile proteins, cyclins, growth factors, oncogenes, integrins, cell surface proteins, etc., their receptors, as well as other receptors. Of particular interest are the UTRs, more particularly the 3'UTRs from mammalian tropomyosin, more particularly the tropomyosin *α*-gene, more particularly the isoforms 2 and 3, troponin I, *α*-actins, such as *α*-cardiac actin and *α*-skeletal actin, vimentin, vinculin, collagen, thrombospondin, VLA antigens, cell adhesion molecules, such as ELAM-1, GMP-140, L-selectin, or other selectins, Peripheral lymph node addressin, mucosal addressin, and the like. The sequences may be associated with the helix-loop-helix (bHLH) transcriptional regulators, cyclins or kinases in controlling cell division. Exemplary sequences of UTRs associated with cell differentiation and/or cell division are set forth in the Experimental section. The genes from which the UTRs are derived will normally be other than housekeeping genes. By housekeeping genes are those genes that are expressed in all or substantially all cell types, are normally consitutive, and may be illustrated by genes associated with metabolism, DNA replication, RNA transcription, and the like

The UTR sequence may be the naturally occurring sequence, fragments thereof, or mutated sequences thereof. Depending on the purpose of the fragments, the size of the fragments may vary widely. The fragments will usually be of at least about 12nt, frequently at least about 20 nt, usually at least about 30 nt, more usually at least about 60 nt and may be 100 nt or more, and will generally not exceed 300nt. The smaller sized fragments up to the full UTR may be used as probes to probe genomic or cDNA libraries for regulatory RNA sequences. Particularly, since the UTR acts as dsRNA,

one can probe with sequences which are complementary to determine whether the UTR is able to form a dsRNA structure. Usually not more than about 90%, more usually not more than about 75%, of the nucleotides of the UTR will be necessary for regulation, generally at least about 35%, where the fragment will be able to modulate cell regulation. or be used as a probe for detecting mRNA or cDNA from a cell being assayed.

C. Assays for detection of regulatory dsRNA

Various techniques can be used to detect the occurrence of dsRNA. By denaturing the candidate RNA in the presence of a known naturally occurring UTR capable of regulating PKR activity, one can determine the rate of dsRNA formation, where one of the sequences is labeled. Various assays are known where one can measure the amount of a label bound to a support, where binding is based on hybridization between the candidate RNA and the labeled RNA. By binding the candidate RNA to the solid support and hybridizing the labeled DNA to the bound candidate RNA, the amount of labeled RNA can be determined as a measure of homology.

One may also screen sequences for motifs associated with regulation of PKR. Of particular interest are UTRs which include at least one occurrence of the sequence TGTANA (where "N" intends any ribonucleotide). Of particular interest is when the indicated sequence is part of a loop of a stem and loop structure of the RNA. Instead of the naturally occurring sequence, mutations may be used. It is found that there is substantial conservation across species for a UTR involved in PKR regulation. In comparing the 3'UTRs of α-tropomyosin among murine and human species, the sequences differ less than about 15% in nucleotides for the sequence associated with PKR regulation. Therefore, some changes may be made in the sequence which do not substantially affect the secondary structure of the sequence, without significantly affecting the activity of the sequence. The changes will usually be fewer than 15 number %, more usually fewer than 10 number %, of substitutions, insertions, and deletions, at other than the regulatory binding site. Based on the experience with viral and synthetic sequences, substantial variation in sequence is permitted without seriously

adversely affecting the regulatory activity, so long as the secondary structure is substantially maintained, particularly the dsRNA structure. Usually, mutations will be fewer than 20%, usually fewer than about 10% of the nucleotides in the sequence. The mutations may involve substitutions and/or deletions. Also of interest is combining fragments from different regulatory UTRs, where the fragments may range from about 10 to 90%, usually 20 to 80% of the wild-type 3'UTR sequence. In referring to the UTR, it is intended to refer to the germline sequence as distinct from the poly(A) processed sequence, although the poly(A) sequence may be present and the UTR may be joined to other sequences, such as the 5' coding region or non-native sequences for a variety of purposes, e.g. hybridization to a sequence, expression, combination of sequences, or the like.

Instead of using the native sequence, one may use the antisense sequence, which may be antisense in either direction, 5'-3' or 3'-5'. As described for the UTR sequence, the antisense or complementary sequence may be the naturally occurring nucleic acids or a synthetic nucleic acid, subject to the modifications indicated above. By use of the antisense sequence, one may inhibit or enhance the effect of the UTR sequence, so as to reverse the effect of the UTR modulation. Thus, one may allow for cell division, while inhibiting cell differentiation. The concentration of the antisense sequence will be at least equal to the concentration of the UTR sequence and usually greater.

Instead of an antisense sequence, one may use a ribozyme sequence. The ribozyme would be chosen to cleave the UTR, so as to disrupt the structure associated with its regulatory activity. By defining the region associated with the regulatory activity, the ribozyme homologous sequence may be chosen to destroy the conformation of the UTR involved with growth and/or differentiation regulation. The ribozyme would have homologous sequences to the UTR, and the necessary catalytic center for cleaving the UTR. See, Sarver, et al., (1990) Science 247, 1222-1225. Usually the ribozyme will have at least about 30nt, more usually at least about 50, and may have 100 or more nucleotides as the sequence homologous to the UTR. By defining conserved regions of the UTRs, the homologous sequences can

be designed to inactivate a plurality of UTRs. For exampl , one could choose to cleave proximal to or in the TGTANA sequence of UTRs having the sequence. Alternatively, one could have a number of different ribozymes having complementary or homologous sequences, so that a number of 3' UTRs would be inactivated, particularly where there is redundancy in function for different 3' UTRs. Other compositions comprise homologous sequences joined to chelated metal ions or other agents which in the presence of cellular oxygen, peroxide or other nucleic acid cleaving agent, serves to cleave the ribonucleic acid when the homologous sequence is bound to the target ribonucleic acid.

The regulatory sequences may be used in a variety of ways. The RNA or DNA sequences, which may be naturally occurring nucleic acids or modified synthetic nucleic acids, may be used *in vivo* or *in vitro* with cells, whereby the sequences are taken up by the cells and modulate the cellular regulation. Thus, the nucleic acid may have the naturally occurring backbone or may have a modified backbone, where heteroatoms, such as oxygen and phosphorus may be substituted by carbon, nitrogen, sulfur, or the like, so as to inhibit degradation of the nucleic acid. Thus, the oxygen may be substituted by methylene, sulfur, or nitrogen, to result in thiophosphates, alkylphosphonates, phosphoramidates, or the like. Alternatively, reduced forms of the phosphorus esters may be used, such as phosphinates. In addition, different sugars may be employed in place of ribose, such as arabinose. Other modifications may also be employed, where the native conformation of the 3'UTR is retained, but the stability to degradation by nucleases is substantially enhanced.

The nucleic acid sequences comprising the unnatural nucleotide need only have one nucleotide to enhance stability or provide some other desired characteristic, usually having two or more, generally at least 50% of the nucleotides being unnatural and frequently the entire nucleic acid sequence. Desirably at least one terminal nucleotide is unnatural. The nucleic acid sequences will be able to compete with the natural sequence for binding to receptors or hybridize to the natural s quence where an antisense sequence is involved.

The nucleic acid sequences which are employed exogenously will usually be purified to at least 50 weight % of the nucleic acids present in the composition, more usually at least 75 weight %, preferably at least 95 weight % and in many situations will be at least substantially pure, being free of any other substances which might interfere with the purpose for which the nucleic acid is being used.

Usually, the nucleic acid will be administered by any convenient means, desirably localized to the target cells. Localization can be achieved by liposomes, having specific markers on the surface for directing the liposome to the target cell, by having injection directly into the tissue containing the target cells, by having a depot associated in spatial proximity with the target cells, specific receptor mediated uptake, or the like. *In vitro*, the medium will include the nucleic acid at the desired concentration. Usually, the concentration will be in the range of about 1ng/ml to 1mg/ml.

The concentration of the nucleic acid sequences will vary, particularly when administered *in vivo*, depending upon the stability of the sequence, its rate of clearance, the environment in which it is acting, the toxicity, if any, the degree and nature of modulation desired, the mode of administration, and the like. Thus, in each application, there will usually be required an empirical determination of the optimum concentration for the desired result.

For *in vivo* administration, where other than a depot is employed, the administration may be in any convenient physiologically acceptable medium, such as water, phosphate-buffered saline, aqueous ethanol, glycerol, mineral oil, and the like. Administration may be as a solution, dispersion, gel, capsule, etc.

The subject compositions can be used by themselves or in conjunction with other compositions for the treatment of neoplasia or other cellular proliferative disease. Thus, the subject UTRs may serve to control cell division of neoplastic cells. The UTRs, either individually or in combination of different UTRs may be applied to the cells as described above. This may be used in conjunction with therapies, such as X-irradiation, cytotoxic agents, such as immunotoxins, alkylating agents, anti-metabolites, antitumor antibiotics, etc. The subject compositions may be used in such techniques

as described in U.S. Patent No. 4,619,913, where the UTR is combined with collagen and injected directly into the lesion.

Instead of adding the UTR, antisense sequence or ribozyme as an exogenous composition, one may modify cells, particularly associated with gene therapy, by introducing a construct capable of transcription of the UTR, the antisense sequence or the ribozyme. The UTR sequence may be by itself or be fused to its endogenous gene or to a sequence other than the sequence of the wild-type gene. By appropriate choice of the promoter, one can provide for inducible or constitutive transcription of the construct. In the case of gene therapy, where one wishes to provide that the cells introduced are not involved in cell division or one wishes to allow for a predetermined level of cell division, which would then be terminated, one may introduce a construct comprising the promoter and the UTR, by itself or with its endogenous gene. By having a threshold level of the UTR being produced in the cell, one can inhibit cell division and allow for progenitor cells to mature into the desired cell type. For example, one could use the UTR with hematopoietic high proliferative potential cells, which are known to be progenitors to the myelomonocytic lineage. By providing for an inducible promoter, one could allow for expansion of the high proliferative potential cells ("HPP cells") until one wished to have mature cells in the myelomonocytic lineage. One could provide for expression of a fusion of the gene encoding G-CSF or GM-CSF and a UTR which modulates cell division. Thus, transcription of the fused mRNA would result in secretion of a protein which would direct the HPP cells into the granulocytic lineage, while stopping cell division of the HPP cells. Alternatively, one may wish to prevent cell division altogether, for example, where one wishes to have the cells cease growing after a predetermined life cycle or one wishes to maintain cells in a particular state where desired product(s) are produced. Thus, one could genetically engineer progenitor cells, including constitutive transcription of the UTR, so as to prevent significant expansion of the progenitor cells, while directing the progenitor cells to cell differentiation and, as desired, maturity. In contrast, one may wish to have a high degree of cell division in a patient, before cell differentiation. In this situation, on can use the antisense

sequence or a ribozyme, which can be produced inducibly, until turned off by an appropriate mechanism. Thus, one could allow for transcription of the antisense or ribozyme sequence, so as to encourage cell division until the cell population had achieved a desired level. One could then use an agent which would turn off the transcription of the antisense or ribozyme sequence, so that the UTR directing cell differentiation could then control the cell cycle. In this manner, one can provide for the introduction of a relatively small number of cells with the assurance that the cells will multiply in the host, until a signal is given, which would then direct the cells to differentiation.

One may also have an interest in maintaining a cell in culture at a primitive or progenitor state. For example, in the expansion of stem cells, e.g. hematopoietic stem cells, one may wish to expand the number of stem cells while preventing differentiation. For bone marrow transplants, it is believed that it is the hematopoietic stem cell which becomes engrafted and serves to repopulate the host. Therefore, during growth of the bone marrow or stem cell enriched cellular composition, one could add anti-sense or ribozyme sequences to the medium to substantially reduce the level of cell differentiation. One may provide compounds which induce the transcription of the regulatory UTRs, such as TGF-β, IGF-I or II, PDGF, or the like.

For cells in culture, depending upon whether differentiation or growth is desired, the amount of the regulatory dsRNA may be varied. At lower concentrations, growth will be inhibited, while at higher concentrations, differentiation may be inhibited. Therefore, one may obtain opposite effects with the same composition, depending on the concentration employed. With the active fragment of the α-tropomyosin 3'UTR, the minimum concentration should be in excess of about 1.0 ng/ml and be less than about 1000 ng/ml, preferably being at least about 2.0 ng/ml and not more than about 600 ng/ml, conveniently in the range of about 10 to 500 ng/ml. These concentrations are consistent with earlier studies using synthetic or viral dsRNAs, that is, derived from other than genomic sources. In vivo, localized concentrations should fall within the above ranges and may be optimized in accordance with the method of adminstration, purpose, stability of the RNA, activity of the RNA, and the like.

In the case of cell mediated gene therapy, where there may be concern for tumorigenicity, by employing a construct which allows for high levels of transcription of the regulatory UTRs, one can substantially reduce the potential for tumorigenicity. In cell populations where one wishes to follow programmed cell death as in apoptosis or cell senescence without concern about cell division, the subject UTRs may be introduced into the cells for high-level constitutive expression. One may also use the subject UTR transcription or exogenous source of UTR to study genes associated with cell differentiation, where subtraction may be employed between a library from cells in which the UTR is transcribed and a library from cells in which it is silent.

In accordance with the subject invention, cells, either normal or neoplastic, can be produced containing cell regulatory amounts of UTR transcripts, as a result of contacting the cells with exogenous UTRs, their complementary sequence (anti-sense), ribozyme directed to the UTR (See, for example, Szostak, Trends in Biochemical Sciences, 1992, 17:89-93; and Bartel and Szostak, Science 1993, 261:1411-8), targetting an endogenous gene comprising the UTR with an enhancer or promoter to activate transcription, introducing an inducible gene which activates transcription of the regulatory UTR, or as a result of introduction of a construct comprising a promoter and the UTR sequence. These cells can be used for investigating regulatory pathways, producing UTRs for use in diagnosis or therapy, controlling cell division, etc. Thus, the cells would inducibly or constitutively produce high levels of UTR RNA, which could then be harvested in accordance with conventional techniques.

D. Diagnosis with UTRs

The subject UTR sequences may also be used in the diagnosis of cell status. Thus, where the UTR is below a predetermined threshold level, this will indicate that the cell may be in a cell division cycle, as contrasted to cell differentiation or vice versa. In addition, the absence of the UTRs may be related to the potential for neoplasia, the initial stages of neoplasia, or the virulence of the neoplastic cell. Alternatively, the enhancement of the UTR

in the cell may be indicative of the effectiveness of therapies or the response of the host to the neoplastic cell or viral infection or viral transformation of cells.

As probes, the sequences may be rendered functional by being joined to a sequence which will hybridize to a sequence bound to a support, have a radioisotope capable of providing a detectable signal, have a label capable of providing a detectable signal, or bind to a surface, e.g. biotin, avidin, fluorescer, enzyme, etc. The nucleic acid may be conjugated in conventional ways to the label, there being a large number of assays using probes described in the literature. See, for example, Sambrook et al., Molecular Cloning: A Laboratory Manual, 2nd ed., 1989, Cold Spring Harbor, NY.

The subject UTRs may also be used for isolating those transcription factors or other regulatory proteins associated with the subject UTR RNAs. Various techniques may be employed for identifying proteins which bind to the subject UTRs. It may be desirable to identify the minimal functional RNA by deletion and/or mutagenesis studies, which are conventionlly employed today. The minimal essential sequence may serve to minimize the number of false positives for binding of the UTR to proteins, as well as identify secondary structure, which may be essential for the binding to the protein. In identifying a protein, for example, an affinity column may be prepared, where the subject UTR RNAs are bound, either covalently or non-covalently, to a gel, solid surface or particles, such as Sepharose, latex, polystyrene, or the like. Ligands on the column may be selected, which specifically bind to a specific RNA conformation. (see, for example, Ellington and Szostak, Nature 1990, 346:818-822.) Cells which are dividing or differentiating may be lysed and the lysate passed through the affinity column or combined with the particles and incubated for sufficient time for complexes to form. The bound proteins may then be isolated in accordance with known ways, purified and sequenced. Alternatively, one may provide for labeled UTR RNAs which are combined with cellular lysates, followed by crosslinking, either chemical or UV, and the resulting composition electrophoresed or chromatographed, whereby complexes between the radioactive UTRs and proteins can be identified. The proteins may also be identified by probing a

λgt11 expression library or other expr ssion library with the UTR RNA. Similarly, the proteins may then be purified and sequenced.

The proteins may be partially or completely sequenced. From the known sequence, one can prepare a family of degenerate probes, which may then be used to screen a cDNA library. Techniques for screening cDNA libraries are well known, see, for example, Sambrook, et al., *supra*. Those clones which bind to the family of probes, may be isolated and undergo a second round of hybridization. Those clones which still are positive for hybridization, may then be sequenced and compared to sequences which have been reported, to determine whether the proteins are known. In addition, one may sequence the cDNA, so that one can identify the total sequence of the protein, where the protein was not totally sequenced.

The subject UTRs may also be used for the preparation of antibodies. Because of the complex secondary and tertiary structures of the subject UTR RNAs, antibodies can be obtained which will be specific for the UTR RNAs. Preparation of antibodies to nucleic acids is described in U.S.Patent No. 4,723,847. These antibodies may then be used in diagnostic assays to identify the presence of the UTRs in cells. Cells in which the UTR is to be identified may be lysed under conditions where nucleases are inhibited, so that the RNA sequences may survive. The antibodies may then be used in a variety of competitive or non-competitive assays, for detection of the presence of the UTR and its amount. Thus, one could combine the sample with a known amount of labeled UTR and a known amount of antibody and determine the amount of labeled UTR which becomes bound as a result of complex formation between the UTR and the antibody. The more UTR present in the cell, the less of the labeled UTR which will become bound.

The antibodies may be monoclonal or polyclonal, being prepared in conventional ways. The UTR may be joined to an immunogen, such as a large protein, e.g. keyhole limpet hemocyanin, HBsAg, or the like, or an immunodominant sequence may be employed to which the host is particularly responsive. Adjuvants may be employed, such as alum, complete or incomplete Freund's adjuvant, or the like. One or mor booster shots may be involved. Various species may be immunized, such as mice,

rats, rabbits, or larger domestic animals for production of the polyclonal antisera. For monoclonal antisera, the spleen may be isolated and the splenocytes fused with an appropriate fusion partner to immortalize the cells. The resulting hybridomas may then be screened for the determination of monoclonal antibodies which are specific for the target UTR.

If desired, the antibodies may be used as immunogens for production of anti-idiotype antibodies, where the conformation of the antibody will compete with the UTR for binding. In this way, the variable region of the anti-idiotypic antibody may be used in place of the UTR for many of the purposes described above. Thus, the anti-idiotypic antibody may be used for complexing with factors which would bind to the UTR, for isolation and analysis. These reagents could have therapeutic application in place of the UTR themselves and be easier to deliver.

The antibodies may be purified as antisera or as monoclonal antibodies, as to specificity, isotype or the like. Preferably, the antibody composition will have at least 50 weight % of antibodies which have the desired characteristic(s), usually at least 75% and may be 95% or greater, where substantial purity can be achieved, paticularly with monoclonal antibodies.

The following examples are offered by way of illustration and not by way of limitation.

EXPERIMENTAL

Materials and Methods

Cell lines, maintenance, fusion, and selection conditions

The C2C12 mouse myoblast line (Blau et al. (1983) *Cell* 32, 1171-1180) derived from C2 (Yaffe and Saxel (1977) *Nature* 270, 725-727) was used: it is diploid and capable of expressing all four of the myogenic helix-loop-helix regulators (Peterson et al. (1990) *Cell* 62, 493-502). Ten subclones of C2C12 cells were plated at low density in high serum medium and ~200 clones of each subclone tested for differentiation potential in low serum medium. From among these, the clone C2F3 was selected for

mutag nesis because 100% of clones contain d myotubes. C2F3 differentiates at low density in low serum medium (2% horse serum). In addition, C2F3 can be maintained as proliferating cells in high serum medium (10% calf serum). Cells were mutagenized by a thirty minute incubation at 37°C with nitrosomethylurea (NMU, Sigma) at doses ranging from 1.25 - 100 μg/ml in phosphate buffered saline. To assess the efficiency of mutagenesis, we monitored the number of $HPRT^-$ mutant clones that grew after exposure to NMU by plating one aliquot of $5x10^5$ mutagenized cells in 10 μg/ml 6-thioguanine (Sigma). To assess the frequency of anchorage independent clones, a similar dose-response curve was generated by plating another aliquot of $5x10^5$ NMU treated cells in 0.34% agar (Difco) according to the procedure described by Bouck and DiMayorca (1982) *Mol. Cell. Biol.* 2, 97-105. Colonies were counted three weeks later. The NMU2 mutant was cloned from soft agar following treatment of C2F3 cells with 12.5 μg/ml NMU and recloned from liquid media to ensure its single cell origin.

All cells were routinely grown in Dulbecco's modified Eagle's medium with high serum (8% or 10% calf serum) in 8% CO_2 at 38°C. The medium was changed every two days. For selection of drug-resistant stable transfectants with SV2 plasmids, the medium was supplemented with G418 (400 μg/ml, Gibco), hygromycin (200 μg/ml, Calbiochem), or puromycin (1 μg/ml, Sigma). For assaying the activation of the muscle promotor constructs in NMU2 cells, the concentration of the appropriate drug was empirically determined for each test cell such that uninduced cells would not survive. NMU2-HCAgpt.3 cells were selected in 7.5 μg/ml mycophenolic acid in media also containing 50 μg/ml xanthine, 7 μg/ml hypoxanthine, 2.5 μg/ml thymidine and 2.2 μg/ml aminopterin (Sigma). The activation of the HCA-puro reporter construct in NMU2-HCApuro.7 cells was determined by selection in 25 μg/ml puromycin. The activation of the MYG-puro reporter construct in NMU2-MYGpuro.7 cells was assessed by selection in 1 μg/ml puromycin.

Fusion of cells to obtain hybrids was performed using a modification of previously reported procedures (Blau et al. (1983) *Cell* 32, 1171-1180;

Chiu and Blau (1984) *Cell* 37, 879-887). Approximately $1x10^5$ cells of each cell type were mixed, rinsed in PBS, and lightly pelleted by centrifugation. The cell pellet was gently resuspended in 250 μl of 50% polyethylene glycol fusion media for one minute, then diluted with DME and rinsed 3 times before plating.

Northern hybridization and probes

Total RNA was isolated by centrifugation of the cell lysates over a cesium chloride cushion (Chirgwin et al. (1979) *Biochemistry* 18, 5294-5299). RNA samples were electrophoresed through 1% agarose-formaldehyde gels for 400 volt-hours, and transferred to Nytran hybridization membranes (MSI) using a vacuum blotter (LKB). The RNA was cross-linked to the membrane by exposure to ultraviolet light (Stratalinker, Stratagene), and hybridized to the probes in a phosphate buffer as previously described (Peterson et al. (1990) *Cell* 62, 493-502). Some membranes were reprobed following stripping of the previous probe in 50% formamide, 5x SSC for 30 minutes at 65°C. Probes for actin, myosin, beta-enolase, MyoD, myogenin, and myf-5 have been previously described (Peterson et al. (1992) *Dev. Biol.* 151, 626-629). The gpt coding region probe was isolated as a 0.6 kb HindIII to EcoRV restriction fragment from the SV2-gpt plasmid. cDNA probes from the clones described herein, including troponin I and tropomyosin (also designated as F1 or Tm) were isolated as the Xho I fragment from the pCDM8 vector (Seed and Aruffo (1987) *Proc. Natl. Acad. Sci. USA* 84, 3365-3369). 25 ng of each DNA was labeled with alpha-^{32}P dCTP (Amersham) using random hexamers as primer (Random Prime Kit, Amersham).

cDNA library construction and transfections

Cells from the muscle tissue of a human fetus at week 12 of gestation were separated into myoblast and non-myoblast (fibroblast) populations based on differential staining with 5.1H11 antibody and fluorescence activated cell sorting (Webster et al. (1988) *Exp. Cell. Res.* 174, 252-265) and cultured as previously reported (Blau and Webster (1981) *Proc. Natl.*

Acad. Sci. USA 78, 5623-5627). The identity of the two populations was confirmed by observing that only the 5.1H11-positive myoblasts were able to fuse and express myosin heavy chain following three days in low serum medium. Poly A+ RNA was isolated from the two cell types after culturing 70% confluent dishes in low serum for 36 hours. The mRNAs were reverse transcribed to cDNA and inserted in random orientation into the pCDM8 mammalian expression vector which utilizes the human cytomegalovirus promoter (Seed and Aruffo (1987) *supra*, Librarian kit, Invitrogen). Each library contained approximately $>1x10^6$ independent colonies with an average insert size of 1.4 kb.

In order to enrich for muscle regulators, single-stranded myoblast cDNAs that hybridized to the biotinylated single-stranded fibroblast cDNA were removed. Single stranded phagemid DNA was prepared from both libraries (random oriented cDNA inserts) using the M13 origin of pCDM8. The fibroblast phagemid library was photobiotinylated and hybridized to the myoblast phagemid library (Subtractor kit, Invitrogen). Myoblast phagemids that hybridized to fibroblast phagemids were removed by phenol extraction and the remainder (subtracted library) were recovered. The subtracted library was transformed into host bacteria MC1061-P3 and plated on agar plates containing tetracycline. Plates were individually harvested 16 hours later and frozen in 40 pools of 500-1000 colonies.

All transfections were performed using the calcium phosphate method of Graham and Van der Eb (1973) *Virology* 52, 452-467. Cells were plated at approximately $1x10^5$ per 60mm dish between 4 and 16 hours prior to transfection with the calcium phosphate precipitate. The medium was replaced eight hours after transfection, and following an additional 24-36 hours of incubation, cells were harvested and replated on two 100mm dishes under selection conditions. Library transfections were performed using 10 μg of plasmid per dish of cells. DNA from 4 subtracted library pools (~4000 cDNAs) was combined and 7 pools of 4000 cDNAs were individually transfected.

Recovery of cDNA from transfectants by PCR

Colonies of cells were isolated using cloning cylinders (Belco) and rinsed 3 times with PBS. The pelleted cells were then lysed by boiling in 100 μl of water, vortexed, centrifuged at 13000xg for 5 minutes, and the supernatant recovered. 25 μl of each supernatant was used as template in the polymerase chain reaction (PCR). Nested PCR was performed using two sets of primers that flank the cDNA insertion site of the vector. Primers used in the first amplification reaction were R1 5'-GAGACCGGAAGCTTCTAGAGATCCCTCGACCTCGAG and L1 5'-TCTAGAGTCGCGGCCGCGACCTGCAGGCGCAGAACT. PCR was performed in 100 μl for 30 cycles of 1 minute at 94°C, 1 minute at 60°C and 2 minutes at 72°C with a final extension of 10 minutes at 72°C using Taq polymerase and conditions recommended by the manufacturer (Perkin-Elmer Cetus). 2 μl of the product was removed and used as template in a second round of amplification using nested primers that correspond to vector sequences located between the first set of primers and the cDNA insertion site (R2 5'-TCGACCTCGAGATCCATTGT and L2 5'-TAGGTATGGAAGATCCCTCG). Reaction conditions for the second round were the same as for the first round.

For directional cloning of the recovered bands into the pCDM8 expression vector, 2 μl of the second reaction was further amplified with primers similar to those used in the second round, except that these primers (R2 Hin and L2 Not) included HindIII and NotI restriction sites added to their respective 5' ends. The amplified products were digested with HindIII and NotI enzymes (BRL) and purified by electroelution from a 2% agarose gel following electrophoresis. Each of the recovered bands was separately ligated into pCDM8 vector that had been restricted with HindIII and NotI enzymes to remove the stuffer fragment.

Test constructs and subcloning of cDNAs

To construct the HCA-gpt reporter plasmid, the 485 bp human α-cardiac actin promoter (EcoRI to HindIII fragment from HCA-CAT, L. Kedes LK339 (Minty and Kedes (1986) Mol. Cell. Biol. 6, 2125-2136;

Hardeman et al. (1988) J. Cell. Biol. 106, 1027-1034) was ligated upstream of the 1.9 kb HindIII to BamHI fragment from SV2-gpt containing the gpt coding sequences, small t splice, and the SV40 polyadenylation signals. The HCA-puro construct was constructed using the 0.9 kb HindIII to BamH1 fragment of pBSpac-Δp (SV2-puro, De La Luna et al. (1988) *Gene* 62, 121-126) encoding puromycin N-acetyl transferase. The MYG-puro reporter plasmid was constructed by 3-fragment ligation using the 184 bp PstI to HindIII mouse myogenin promoter (Edmondson et al. (1992) Mol. Cell Biol. 12, 3665-3667), the 0.9 kb puromycin N-acetyl transferase coding region and the 2.7 kb PstI to HindIII fragment of pUC19 vector.

The tropomyosin 5'-coding sequence was isolated from the library using PCR with a specific primer that encompasses the termination codon of tropomyosin (5'-TCACATGTTGTTTAACTCCAGTAA) and a primer corresponding to cDNA-flanking library vector sequence (R2 Hin). UTRs of oct 1 and jun B were derived using published sequences (Sturm et al. (1988) *Genes and Dev.* 2, 1582-1599; Nomura et al. (1990) *Nucleic Acids Res.* 18, 3047-3048). Template DNA was generously provided by G.R. Crabtree, Stanford University, Palo Alto CA. Primer pairs for the two sequences were 5'-CCCTCGAGCTGGGCAGAGCTGG-3',5'-CCCCCCGCGGCCGCTTTATTA CAAAATTAGG-3', and 5'-GGCTCGAGAACGTCCCCTGCCCCT-3', 5'-CCCCGCGGCCGCTTAAATAGATTCAATAAAAAGAA-3' respectively. Histone H3.3 UTR (Wells et al. (1987) *Nucleic Acids Res.* 15, 2871-2889) was obtained by PCR from our human myoblast library using a flanking vector sequence primer (R2 Hin) and a specific primer 5'-TTGCGGCCGCTAGGTATGGAAGATCCCTCG-3'. The primers were designed with appropriate restriction sites to allow directional cloning of the PCR-generated fragment into HindIII and NotI-digested pCDM8 vector. The G12 cDNA (troponin I) 5'-coding and 3'-UTR segments were obtained using PCR and primers that allow each portion to be separately amplified and ligated into pCDM8 in the sense orientation. To obtain promoterless F1, the 0.6 kb CMV promoter sequences were deleted from pCDM8 by restriction digestion with MluI and SacI, blunt end formation with T4 DNA polymerase (Dupont), followed by ligation with T4 DNA ligase. The UTRs derived from

K1, histone H3.3, D3, H3 and G12 contained poly A tails, whereas F1, oct 1, and jun B did not. The constructs were confirmed by sequencing. All plasmids used were purified by double banding on cesium chloride gradients. Levels of expression of UTRs in stably transfected cells were in the range of the levels expressed by endogenous genes.

Immunostaining

Myogenin expression was assayed using monoclonal supernate from hybridoma line 1F5D7 (a gift from W. Wright; Lassar et al. (1991) *Cell* 66, 305-315). Colonies were chosen by visual inspection, circled, and analyzed at the single cell level for nuclear myogenin expression by immunocytochemistry. Cells were rinsed with PBS and fixed with 2% formaldehyde for 10 minutes at room temperature and permeabilized with 1% NP40 (BDH Chemicals) in PBS for an additional 10 minutes. Rinsing and incubation solutions used thereafter contained 2% horse serum and 0.1% NP40. Permeabilized cells were incubated in rinsing solution for 2 minutes, incubated in the monoclonal supernatant for 1 hour, rinsed twice and incubated with biotinylated anti mouse-IgG secondary antibody (1:200, Vector Laboratories) for 30 minutes. The staining was developed using Vectastain horseradish peroxidase kit (Vector Laboratories) using the manufacturers recommended conditions. Colonies were examined and counted using bright field optics at 100x magnification on a Leitz Labovert microscope and the frequency of the colonies containing myogenin-positive nuclei were determined. The sensitivity of this assay was confirmed by introducing a myogenin cDNA expression vector into the test cell; greater than 90% of stably transfected clones expressed detectable myogenin protein in their nuclei. For each cDNA tested, a minimum of 150 colonies was scored, the cumulative results of at least two independent transfection experiments with two independent preparations of the plasmids. As is commonly the case for myogenic colonies, positive cells were generally detected in the denser central region of the colonies: only rare positive cells were found at the edges.

Immunostaining for sarcomeric actin and N-CAM was performed using monoclonal antibody B4 (Lessard (1988) *J. Cell Motil. Cytoskel.* 10, 349-362) and anti-mouse N-CAM (Chemicon) at 10 μg/ml. Staining was visualized using FITC-conjugated goat anti-mouse IgG (actin) or goat anti- rat IgG (N-CAM).

Statistics

In all cases, the error indicates the standard error of the proportion calculated from the standard binomial equation: $\sqrt{(pq/n)}$. Where error bars do not overlap, differences are significant at the 0.05 level, using a two-sample Student's t test.

Stable Transfectants for Assay of Transformation

NMU2 Tm Clone 1 and control clones 2, 3, and 4 were derived by stable transfection of NMU2-HCAgpt.3 with the tropomyosin (F1) 3'UTR. All other NMU2 transfectants were obtained by transfecting the parental NMU2 cell line. The cotransfected selectable marker for NMU2 Tm Clone 1 was SV2-hygro. The other transfectant HT1080 was cotransfected with SV2-puro. Transfections were performed using the calcium phosphate method of Graham and Van der Eb, 1973, *supra*. Cells were plated at approximately $1x10^5$ per 60mm dish between 4 and 16 hours prior to transfection with the DNA precipitate. 2 μg of selectable marker plasmid (SV2-hygro or SV2-puro) and 8 μg of cDNA expression plasmid (in pCDM8 vector) were added to each dish. The medium was replaced eight hours later, and following an additional 24-36 hours of incubation, cells were harvested and replated under selection conditions. Clones of stable transfectants were isolated after 7 to 10 days using glass cloning cylinders.

Anchorage Independence Assay

To assess anchorage independent growth, cells were suspended in agar (Difco) according to the procedure described by Bouck and DiMayorca, 1982. *supra*. Each cell was tested in duplicate. Harvested cells were counted using a hemacytometer and 200 cells wer plated in duplicate

dishes of liquid medium to assess viability. Cloning efficiency of each cell type in medium containing agar was corrected for cloning effici ncy in liquid medium in order to normalize for viability of the cells at the time of plating. For determining cloning efficiency in agar, one to ten thousand cells were plated in 1.5 ml of media containing 0.34% agar. This suspension was layered over 5ml of a base layer of solidified media containing 0.5% agar. NMU2 cells were tested in DME with 8% calf serum. After each week of culture, 1 ml of the appropriate liquid medium was layered over the agar cultures to prevent desiccation. The size and number of colonies were counted after three weeks.

Tumorigenicity Assay

To assay for tumor formation in vivo, cultures of cells to be tested were harvested, rinsed with PBS, and resuspended in PBS. An inoculum of $1x10^5$ cells in 50 μl of PBS was injected into each site. Cells were injected subcutaneously at two sites in the flanks of athymic nude mice. In independent experiments, either all male or all female mice were used and the animals ranged in age from 7 to 10 weeks. Tumor formation was assessed twice weekly for 10 weeks. Data show the time when tumors were first detected. In all cases shown, tumors continued to grow until the animal was sacrificed. When necessary, tumors of approximately 5mm diameter were removed after sacrificing the animals. Tumor cells were recovered by extensively mincing the tumor tissue. The minced tissue was cultured, and adherent cells were isolated.

Northern and Southern Blot Analysis

Total RNA was isolated by centrifugation of the cell lysates over a cesium chloride cushion (Chirgwin et al., 1979, *supra*). RNA samples were electrophoresed through 1% agarose-formaldehyde gels for 400 volt-hours, and transferred to Nytran hybridization membranes (MSI) using a vacuum blotter (LKB). The RNA was cross-linked to the membrane by exposure to ultraviolet light (Stratalinker, Stratagene), and hybridized to the probes in a

phosphate buffer as pr viously described (Peterson et al. (1990) Cell 62, 493-502).

For Southern analysis, DNA was isolated from cells in culture according to the method of Fukui et al. (1982) Proc. Natl. Acad. Sci. USA 82, 5954), and digested with HindIII. 10 μg of DNA was electrophoresed through a 1% agarose gel and transferred to Nytran membrane. Southern blots were prehybrized for 8 hrs in 5x SSC, 5x Denhardt's, 0.1% SDS and 100 μg/ml sheared salmon sperm DNA.

cDNA probes were isolated from the pCDM8 vector by restriction-enzyme digestion and separation of the Xho I fragment on a 2% low-melting temperature agarose gel. Five μl of the gel slice containing approximately 25 ng of DNA were labeled with alpha-^{32}P dCTP (Amersham) using random hexamers as primer (Multiprime Kit, Amersham).

Histology of muscle implanted with β-gal-expressing cells

To analyze the fate of the cells histologically, the *lacZ* gene was introduced by retroviral infection into both NMU2 cells and NMU2 cells stably transfected with the Tm 3'UTR. The MFG retrovirus (provided by Paul Robbins, University of Pittsburgh) drives expression of *lacZ* from the MoMuLV long terminal repeat. Virus-containing supernatant was obtained from confluent amphotropic packaging cells. Repeated rounds of retroviral infection were carried out until more than 95% of NMU2 cells and stable transfectants expressed β-galactosidase, as determined by histochemical assay of enzyme activity (Hughes and Blau, (1992) Cell 68, 659-671).

In nude mice 5-8 weeks of age, $1x10^5$ to 2×10^5 cells were implanted into the interior of tibialis anterior muscles, directly visualized by making an incision in the overlying skin. This method of cell delivery overcomes the problems associated with blind injection, in which cells are often implanted between two muscles rather than within a given muscle. In general, myoblasts injected between muscles do not have access to muscle fibers. Mice were sacrificed at different times up to 3 weeks after implantation. The muscles were removed and frozen, and serial sections were obtained by cryostat for histochemical analysis. The fate of implanted cells was

determined by staining fixed 30 μm sections for β-galactosidase activity (Hughes and Blau (1992), *supra*). Adjacent 10 μm sections were stained with hematoxylin and eosin for histological analysis of tissue morphology and architecture in the same regions in which injected β-galactosidase-labeled cells resided.

RESULTS

Isolation of Differentiation-defective Myogenic Mutants

The genetic approach requires that introduction of a single cDNA lead to a detectable change in differentiation specific gene expression, a property that is difficult to predict for any given cell type. Differentiation-defective myogenic mutants, cells that had expressed the genes of interest prior to mutagenesis were generated. For this purpose, a stable subclone of C2C12 wild-type myoblasts was isolated and mutagenesis was carried out at low dose to minimize the chance that more than one mutation occurred in a myogenic pathway. In addition, a simple selection procedure was used to identify rare mutants and methods were devised for determining whether a mutant was defective in a regulatory step that could be complemented by regulators present in wild-type myoblasts. This strategy for the production and characterization of mutants is described in detail below.

Prior to mutagenesis, a stable subclone of the C2C12 mouse myogenic cell line was isolated that generated spontaneous differentiation-defective variants at low frequency. The C2F3 subclone was selected for mutagenesis because it rarely gave rise spontaneously to differentiation-defective variants and had the additional attractive feature that it generated anchorage independent colonies at remarkably low frequency, $<1x10^{-5}$, as compared to $5x10^{-2}$ for the C2C12 parent, allowing selection of rare mutants by growth in soft agar.

For mutagenesis, nitrosomethylurea (NMU) was used, a point mutagen for which the mutation efficiency at different doses has been extensively documented (Richardson et al. (1987) *Proc. Natl. Acad. Sci. USA* 84, 344-348; Bouck and DiMayorca (1982) *Mol. Cell Biol.* 2, 97-105). Aliquots of C2F3 cells were incubated in different concentrations of NMU for

30 minutes. Viability, an indirect measure of the number of mutations, was assessed by comparing the plating effici ncy of NMU treated and mock (PBS only) treated C2F3 cells. The mutation rate was determined based on the dose-response curve for inactivating a single well-defined gene, the gene encoding hypoxanthine phosphoribosyl transferase (HPRT) on the X-chromosome. A similar dose-response curve was generated for the production of anchorage independent C2F3 mutants by assaying growth in soft agar. The slope for the production of anchorage independent mutants was somewhat greater than that for HPRT mutants. This finding suggests that more than one mutation can lead to the anchorage-independent phenotype. Based on these data, all myogenic mutants were selected after exposure to 12.5 μg/ml NMU, a low dose at which mortality was negligible and the frequency of generating mutants approximated 10^{-5}.

Growth in soft agar proved to be an effective method of selecting for differentiation-defective myogenic mutants. All of the mutants that grew in soft agar failed to form myotubes in liquid tissue culture media. The mutants expressed different combinations of muscle-specific structural genes and regulatory genes of the myogenic helix-loop-helix (bHLH) family. This finding further suggests that disruption of several different pathways could lead to the anchorage independent growth of myoblasts, in agreement with the dose-response curve.

Characterization of the NMU2 Mutant

NMU2 was isolated as a clone that grew in soft agar and it was recloned in liquid medium to ensure its single-cell origin. Myotube formation was not detected when cells were grown at high density either in 10% calf serum or 2% horse serum, conditions under which C2F3 differentiates. The doubling time for NMU2 in 10% calf serum is 12-14 hours, similar to that of the parental C2F3. NMU2 cells are more refractile and less adherent than C2F3 cells. Total RNA was isolated from cells grown to high density (approximately $3x10^6$ per 100mm dish) in medium containing 10% calf serum. 10 μg of RNA from each cell type was electrophoresed, blotted onto a hybridization membrane, and probed sequentially for th helix-loop-helix

regulators, Myf-5, myogenin, and MyoD, for sarcomeric myosin heavy chain (MyHC), actins, troponin I, and tropomyosins and for the muscle-specific isoform of β-enolase. The NMU2 mutant fails to express mRNAs encoding myogenin, myosin heavy chain, sarcomeric α-actins, troponin I and at least one isoform of tropomyosin. However, NMU2 expresses muscle-specific enolase, MyoD, Myf-5, and non-muscle beta and gamma actins. In addition, nuclear MyoD and cell surface N-CAM proteins are detected in nearly 100% of the cells by immuno-cytochemistry. Thus, it appears that a subset of steps in the myogenic regulatory network have been disrupted in the NMU2 mutant, demonstrating that the control of certain pathways can be dissociated. The finding that multiple muscle products ceased to be expressed in NMU2 suggested that mutagenesis affected a regulatory step and not the structural genes themselves, a point confirmed by the experiments described below.

NMU2 Lacks Activators of the α-Cardiac Actin Promoter

To determine whether the defect in NMU2 involves regulators of myogenic genes, we examined the activity of the α-cardiac actin promoter. Both α-cardiac and α-skeletal actins are expressed in C2F3 cells, but neither one of these two sarcomeric actins accumulates in the NMU2 mutant either at the level of RNA detected by Northern analysis or protein detected by immunofluorescence. To analyze the expression of the α-cardiac actin promoter in NMU2 cells, we used the well characterized 485 bp promoter fragment of the human α-cardiac actin (HCA) gene (Minty and Kedes (1986) *Mol. Cell. Biol.* 6, 2125-2136; Gustafson and Kedes (1989) *ibid* 9, 3269-3283). Stable transfection experiments were performed using a construct (HCA-gpt) in which the human α-cardiac actin promoter drives the expression of bacterial guanine-xanthine phosphoribosyl transferase which confers resistance to mycophenolic acid. The expression of this construct was compared in NMU2 and C2F3 to determine the relative level of promoter activity. As a control for potential differences in reporter gene activity in the two cell types, each was co-transfected with the plasmids SV2-gpt and SV2-neo. C lls from each transfected dish w re divided and grown either in

mycophenolic acid to select for gpt expression or in G418 to select for neo expression. Since the transfection efficiency of the two cell types differed, SV2-neo was used to correct for transfection efficiency in all co-transfections with SV2-gpt and HCA-gpt. SV2-gpt was expressed equally well in both NMU2 and C2F3 cells, whereas HCA-gpt was expressed efficiently only in C2F3 cells. As shown in Table 1, the markedly reduced activity of the *α*-cardiac actin promoter in NMU2 as compared to C2F3 cells indicates that NMU2 lacks the *trans*-acting factors necessary for expression of this promoter.

TABLE 1

Relative Activity of HCA Promoter in C2F3 and NMU2 cells

Co-transfected Plasmids	Colonies/Transfection*					
	Wild-type (C2F3)			Mutant (NMU2)		
	gpt^+	neo^r	gpt^+/neo^r	gpt^+	neo^r	gpt^+neo^r
SV2-gpt + SV2-neo	21	18	1.2	40	52	0.75
HCA-gpt + SV2-neo	29	21	1.4	1	38	0.03
Relative HCA activity#:			1.2			0.04

NOTES TO TABLE 1:

* $1x10^5$ cells were cotransfected with 5 μg of each plasmid. Cells from each of two duplicate transfections were plated in two dishes. One dish was selected in media containing mycophenolic acid (gpt^+) and the second dish selected in medium containing G418 (neo^r). Colonies were counted 14 days later. The number of colonies represents an average of duplicate transfections.

\# Number of gpt^+ colonies in each cotransfection was corrected for the number of neo^r colonies (gpt^+/neo^r). The corrected values for the HCA-gpt plasmid were then compared with the corrected values for the SV2-gpt plasmid in each cell type.

Compl m ntati n of NMU2 and Activation of th *α*-Cardiac Actin Prom ter in Hybrids

To determine that the defect in NMU2 could be complemented by regulators p.ovided *in trans* by a wild-type myogenic cell, fusion experiments were performed. To assay for regulators of the *α*-cardiac actin promoter, we used NMU2 cells stably transfected with a reporter gene construct. NMU2 cells were co-transfected with HCA-gpt and SV2-neo and selected for G418 resistance. Suitable test cells for genetic complementation were identified by their ability to express the HCA-gpt construct after fusion with C2F3 cells in hybrids. Fewer than 3% of the NMU2 transfectants expressed the HCA-gpt construct constitutively. Five NMU2 transfectants that did not express HCA-gpt were fused with C2F3 using polyethylene glycol and selected in media containing mycophenolic acid. We produced hybrids, not heterokaryons, since ultimately complementation by cDNAs had to be performed under conditions that allowed for growth and drug selection. Survival of the hybrids depended on the activation of the HCA-gpt construct in NMU2 in response to *trans*-acting factors provided by C2F3. Three clones did not express gpt after fusion with C2F3. Two clones gave rise to mycophenolic acid resistant (gpt^+) hybrids when fused to C2F3 but not when fused to themselves (homohybrid control). These results demonstrate the necessity of identifying clones for complementation by cell fusion. In addition, they confirm previous findings showing that clones differ in their ability to activate stably transfected constructs due presumably to a combination of copy number of the construct and the nature of the DNA flanking the site of integration (Hardeman et al. (1988) *J. Cell. Biol.* 106, 1027-1034); Spear and Tilghman, 1990, *supra*; Wu et al., 1991, *supra*). The clone, NMU2-HCAgpt.3, which exhibited vigorous growth when fused in a hybrid, was used as a test cell for isolating activators of the silent 485 bp *α*-cardiac actin promoter.

Complementation of NMU2 by cDNAs

To minimize differences, cDNA libraries were constructed from poly A + RNA isolated from primary myoblasts or fibroblasts derived from the

same human muscle tissue. Myoblasts that had just initiat d differentiation following a 36 hour exposure to low s rum medium w re used in order to enrich for regulators of myogenesis. cDNAs were cloned in random orientation into the pCDM8 vector that utilizes the human cytomegalovirus (CMV) promoter, which is efficiently expressed in mammalian cells (Seed and Aruffo, 1987, *supra*). Each library contained approximately $1x10^6$ cDNAs with an average insert size of 1.4 kb. The myoblast library was subtracted with the fibroblast library to enrich for muscle-specific cDNAs. To determine whether intact cDNAs were present in the library and could be expressed in mammalian cells, we isolated actin cDNAs from the library by colony hybridization with an actin probe. Seven of sixteen hybridizing clones had cDNA inserts that appeared full length (1.6 - 1.8 kb) on Southern blots and four of these expressed actin protein as determined by immunofluorescence following transfection into COS cells.

Both subtracted and unsubtracted libraries were introduced into the NMU2-HCAgpt.3 test cells and the cells selected in media containing mycophenolic acid to assay for the activation of the HCA promoter. No colonies were detected in ten control dishes of cells transfected either with salmon sperm DNA or library vector plasmid. Only 3 colonies were detected in a total of twenty dishes of cells transfected with the myoblast cDNA library. By contrast, an average of 15 colonies/dish was obtained when cells were transfected with different pools of the subtracted library. Thus, the subtracted library appeared to be enriched 100-fold for regulators capable of activating the HCA promoter and was used in all further experiments.

To confirm the presence of library DNA in the activated cells, colonies were picked, expanded to approximately $5x10^6$ cells, and DNA isolated Southern analysis. The majority of the clones demonstrated a retarded growth rate relative to the parental NMU2-HCAgpt.3 cells, even when selection in mycophenolic acid was eliminated. In addition, Southern analysis revealed that many of the expanded clones contained fewer than one copy per genome of library vector sequences, which could have been due to loss of transfected DNA during growth of the clones. To overcome the problem of loss of differentiation inducing cDNAs from th activated clones, colonies

were individually harvested when only 50 to 100 cells had accumulated, two to three weeks after the initiation of growth in selective media. Each gpt^+ clone was lysed and nested PCR performed with primers to flanking vector sequences to amplify cDNA inserts. Amplified DNA was detected in 94% of gpt^+ clones, suggesting that they contained library sequences. No DNA was amplified from mock lysates prepared from regions of the transfected dish that did not contain cells. Between one to eleven amplified bands could be detected in the size range of 0.2 to 1.2 kb, although, in general, the products of smaller size appeared to be preferentially amplified.

NMU2-HCAgpt.3 test cells were transfected with 10 μg of DNA from independent pools of the cDNA expression library. Complemented gpt^+ clones were selected in media containing mycophenolic acid. Individual colonies were trypsinized and harvested in cloning rings, the cells rinsed with PBS, and lysed in 100 μl of water. A mock lysate was prepared following a mock harvest from a region of the transfected dish that did not contain cells. Lysates were used as PCR templates to recover cDNAs flanked by library vector sequences. PCR-generated bands recovered from activated clones were visualized by electrophoresis on 2% agarose gels stained with ethidium bromide. Molecular weight markers were employed and bands found to have activity in subsequent experiments were indicated as D3, F1, G12, H3.

To determine which of the amplified products corresponded to a cDNA capable of activating the HCA-gpt construct, PCR amplified DNAs were isolated from agarose gels and tested individually. NMU2-HCAgpt.3 cells were co-transfected with 2 μg SV2-hygro and 8 μg of each of the complementing cDNA expression constructs, the negative control (K1), pCDM8 (Vector), or sheared salmon sperm DNA (Mock). The DNA from each of 50 bands was directionally cloned into the pCDM8 vector and pools of cDNAs co-transfected together with SV2-hygro into the NMU2-HCAgpt.3 test cell. Stable transfectants were grown in medium containing hygromycin, but without mycophenolic acid. Thus, in this experiment there was no selective pressure to express gpt. Between 50 and 400 stable colonies were obtained for each pool of cDNAs and combin d for Northern analysis. cDNAs in positive pools induced significant gpt expression and were retested

individually. Transfection of four single cDNAs repeatedly led to activation of the r porter gene. The total amount of hybridizing signal in each lane reflected the extent of gpt RNA induction by each of the four activating cDNAs. These cDNAs were designated D3, F1, G12, and H3 based on the library pool from which they were derived. No gpt mRNA was detected in negative controls transfected with vector or with salmon sperm DNA (mock) or in cells that received cDNA from amplified bands that did not complement the test cell, such as K1.

cDNAs That Activate HCA-gpt Also Activate HCA-puro and MYG-puro

To determine whether the activation of the HCA promoter by the four cDNAs was dependent on the choice of the reporter gene or on the particular site of integration of the HCA-gpt construct in NMU2-HCAgpt.3, a second test cell was created. The HCA promoter was linked to a different reporter gene, the bacterial puromycin N-acetyl transferase (puro) gene which confers resistance to puromycin. Stable clones were selected by co-transfection with SV2-neo and selection in medium containing G418. In order to identify a test cell in which expression of the construct could be induced when exposed to regulators provided *in trans,* six randomly chosen G418-resistant clones were analyzed. These clones were fused to form homohybrids and heterohybrids with C2F3 as in the case of the first test cell, except that they were selected for growth in media containing puromycin. Of the six clones tested, one expressed the construct constitutively, five did not express the construct, and two of the five could be induced to express it only in heterohybrids. A clone that exhibited vigorous growth in the presence of puromycin as a hybrid fused to C2F3, but failed to grow when fused to itself, was selected as the second test cell and designated as NMU2-HCApuro.7.

To determine whether the activation of the *α*-cardiac actin promoter observed with the four cDNAs was specific to that promoter or had a more general effect on myogenic genes in NMU2, a third test cell was created. In this test cell, we used a different reporter construct, a 184 bp fragment of the mouse myogenin promoter (Edmondson et al., 1992, *supra*) linked to the gene encoding resistance to puromycin (MYG-puro). Myogenin is a member

of the helix-loop-helix family of myogenic regulators and, like cardiac actin, is not expressed in NMU2. Stable transfectants were produced with this construct and a test cell (NMU2-MYGpuro.1) identified in which the reporter construct could be activated in hybrids, as described for the other test cells. Homohybrids of the test cell fused to itself did not activate the construct.

All four cDNAs capable of activating the HCA-gpt construct in the first test cell were also capable of activating the HCA-puro and MYG-puro constructs in the two new test cells (Table 2A). Transfectants were selected in medium containing puromycin, at concentrations empirically determined to kill the untransfected test cells. As a positive control, an expression vector encoding the myogenin protein was transfected into each test cell. The bHLH protein myogenin is a transcription factor that is absent from NMU2 cells (Table 2B) and known to bind directly to consensus elements within both the cardiac actin and myogenin promoters (Sartorelli et al. (1990), *Genes Dev.* 4, 1811-1822; Edmondson et al., 1992, *supra)*.

TABLE 2

cDNAs Activate Muscle-Specific Promoters

A. Activation f α-cardiac actin (HCA) and myogenin (MYG) promoters in two t st cells

Transfection*	Puromycin^r colonies/Transfection	
Construct in NMU2 Test Cells		
	HCA-puro	MYG-puro
Control cDNAs:		
Myogenin	12	49
E47	0	0
K1	0	0
Complementing cDNAs:		
D3	12	12
F1	7	17
G12	20	26
H3	15	26

B. Activation of endogenous myogenin gene

Transfection#	Myogenin Positive Colonies (%)
Control cDNAs:	
E47	1 ± 0
K1	1 ± 0
Complementing cDNAs:	
D3	14 ± 1
F1	29 ± 5
G12	25 ± 3
H3	11 ± 2

NOTES TO TABLE 2:

* $1x10^5$ test cells (NMU2-HCApuro.7 and NMU2-MYGpuro.1) were transfected with 10 μg of plasmid. NMU2-HCApuro.1 cells wer selected in media containing 25 μg/ml puromycin and

NMU2-MYGpuro.1 cells were selected in media containing 1 μg/ml puromycin. Colonies were counted following 10 days of selection. Each cDNA was tested in two or more transfection experiments of which one representative is shown. All cDNAs use the CMV promoter for expression except for the myogenin cDNA which uses the MSV promoter.

NMU2-HCAgpt.3 cells were cotransfected with 8 μg of cDNA plasmid and 2 μg of SV2-puro plasmid. Puromycin-resistant colonies were grown under selection with 1 μg/ml puromycin for 7 days and then maintained without puromycin in medium containing 0.5% horse serum for an additional 2 days. The frequency of myogenin-positive colonies was determined following immunostaining for myogenin protein and is expressed ± the standard error of the proportion. For each cDNA tested, a minimum of 150 colonies was scored, the cumulative results of three different transfection experiments with two independent preparations of the plasmids. Colonies that contained more than two cells with definitive nuclear myogenin staining were scored as positive. Positive cells were generally detected in the denser central region of the colonies.

As a negative control for non-muscle specific gene activation via bHLH proteins, an expression vector encoding E47, a ubiquitous bHLH (Schlissel et al. (1991) *Genes Dev.* 5, 1367-1376) was introduced into each test cell. As a second negative control, one of the PCR amplified cDNAs which had not activated HCA-gpt when introduced alone (e.g. K1 cDNA) was transfected. Neither E47 nor K1 activated the promoter constructs in test cells. These negative controls provide evidence that the reversion frequency is below the level of detection in this assay.

The activation of three different constructs in three different test cells strongly suggests that the observed effect is not dependent on th site of integration of the reporter gene construct. In addition, the results cannot be

ascribed to a peculiarity specific to the metabolism of the report r gene itself, since similar results were obtained for two different bact rial genes that conferred resistance to mycophenolic acid and puromycin. Furthermore, since two different muscle-specific promoters were activated, it appeared possible that the cDNAs could have a broader effect on the differentiated state of NMU2.

cDNAs Activate the Endogenous Myogenin Gene

We examined the potential of the four cDNAs to activate the expression of the endogenous myogenin and sarcomeric actin genes in NMU2 cells. The activation observed in the three test cells described above involved only 184 bp or 485 bp of the myogenin and cardiac actin promoters, respectively. To determine whether the resident endogenous promoters could also be activated, we used the NMU-HCAgpt.3 test cell. The cells were co-transfected with the SV2-puro plasmid and expression vectors containing one of the four activating cDNAs, or the K1 or E47 cDNA negative controls. Clones were selected for uptake of DNA, not for expression of the reporter gpt gene, by growth in the presence of puromycin. Expression of the endogenous myogenin gene was assayed in individual colonies by immunocytochemistry using an antibody to myogenin. The frequency of positive colonies observed with the four complementing cDNAs ranged from 11% for H3 to 29% for F1 in this assay, values well above those observed for the negative controls E47 and K1 which both approximated 1% (Table 2B). These results show that the endogenous myogenin gene can be activated upon expression of the four cDNAs in NMU2 cells.

The activation of endogenous myogenin determined *in situ* by immunocytochemical analysis of individual clones was corroborated by Northern analysis. For this purpose, the test cell NMU2-MYGpuro.1 was co-transfected with the F1 cDNA and SV2-hygro and grown in selective media containing hygromycin. NMU2-MYGpuro.1 cells were co-transfected with 1 μg SV2-hygro and 9 μg of the F1 cDNA and selected for hygromycin resistance. Individual clones were expanded and total RNA was is lated and

analyzed by Northern for the xpression of F1 and myogenin transcripts. NMU2-HCAgpt.3 cells or C2F3 cells were co-transfected with 2 μg of SV2-puro and 8 μg of each cDNA expression construct and selected in 1 μg/ml puromycin for 7 days in medium containing 10% calf serum. The resulting NMU2-HCAgpt.3 colonies were incubated for an additional 2 days in medium containing 0.5% horse serum. Both sets of stable transfectants were fixed, and stained with antibody specific to the myogenin protein. Thus, stable transfectants were selected for uptake of DNA, not for their expression of the reporter gene constructs. RNA was isolated and probed first for expression of the F1 cDNA and then for expression of myogenin. The frequency of colonies of average size (25-250 cells) that contained greater than five cells with definitive nuclear staining was determined by microscopic analysis. Many of the colonies transfected with TnI or Tm UTRs had >25 positive nuclei usually located in the central dense region of the colony. Each determination was based on the cumulative results for two separate experiments with the standard error of the proportion. An average of 230 colonies was analyzed for each UTR transfection.

A total of six clones were tested, 4 of which expressed F1 and induced myogenin mRNA. Two clones did not express F1 and, like the untransfected control, did not induce myogenin expression. To test whether transcription of the F1 cDNA is required for myogenin activation, all of the CMV promoter sequences were deleted from the F1 expression plasmid. Two types of colonies were obtained following the transfection of the promoterless construct, those that expressed a larger transcript which hybridized with the F1 probe and those that had no detectable F1 transcript. Myogenin mRNA was induced only when F1 transcripts were detected. These data suggest that F1 expression is required for myogenin activation.

By contrast with myogenin, endogenous sarcomeric actin transcripts or protein were not detected upon expression of any of the four cDNAs. Following transfection with each cDNA, approximately 30 individual clones were scored *in situ* by immunocytochemistry using an antibody specific for sarcomeric actin (L ssard, 1988, *supra*). None exhibited staining abov background. In addition, α-cardiac actin RNA was not detected by Northern

analysis of the cDNA transfectants. These findings suggest that sequences other than the 485 bp promoter in the endogenous α-cardiac actin gene mediate its repression in NMU2 cells.

Characterization of complementing cDNAs

To determine the nature of the complementing cDNAs, they were sequenced. Sequences of D3, F1 and G12 are identical to human sequences for α-cardiac actin, tropomyosin and skeletal muscle troponin I, respectively (Table 3).

TABLE 3

Identification of cDNAs

cDNA	Size	Identity	Reference
Complementing cDNAs:			
D3	0.7 kb	human alpha cardiac actin	Hamada et al. (1982) *Proc. Natl. Acad. Sci. USA* 79, 5901-5905
F1	0.3 kb	human tropomyosin	Lin and Leavitt (1988) *Mol. Cell. Biol.* 8, 160-168
G12	1.0 kb	Human muscle troponin I	Wade et al. (1990) *Genomics* 7, 346-357
H3	0.3 kb	unknown	
Control cDNAs:			
K1	0.7 kb	human mitochondrial ATPase 6	Anderson et al. (1981) *Nature* 290, 457-465
jun B	0.4 kb	human jun B	Nomura et al. (1990) *Nucl. Acids Res.* 18, 3047-3048
oct 1	0.2 kb	human octamer binding protein	Sturm et al. (1988) *Genes and Dev.* 2, 1582-1599
H3.3	0.5 kb	human histone H3.3	Wells et al. (1987) *Nucl. Acids Res.* 15, 2871-2889

D3 sequence (human *α*-cardiac actin UTR)

CGCAAATGCTTCTAAGATGCCTTCTCTCTCCATCTACCTTCCAGTCAGGA
TGACGGTATTATGCTTCTTGGAGTCTTCCAAACCACCTTCCCTCATCTTT
CATCAATCATTGTACAGTTTGTTTACACACGTGCAATTTGTTTGTGCTTCT
AATATTTATTGCTTTCATAAATAAACCAGACCAGGACTTGCAACCTAA (The sequence differs by two bases from the sequence reported by Hamada et al. (1982) *supra*)

F1 (human tropomyosin UTR) sequence

CAAAAGGTCCCCCTGTGGTCTTTTGTGTCAACATTGTACAATGTAGAACTC
TGTCCAACACTAATTTATTTTGTCTTGAGTTTTACTACAAGATGAGACTAT
GGATCCCGCATGCCTGAATTCACTAAAGCCAAGGGTCTGTAAGCCACGCTG
CTCTTCTGAGACTTCCAATCAGAAAGGAATGGAAGTCTCAGAGAGCCAGC
GTGGCTTACAG (partial sequence of 3'UTR)

G12 (human troponin I UTR) sequence

CTTCTTCCCTCCAGCCTGCAATGCCCTCCTCTGGAACTGGGATTAAACAGA
TACCCAAGAGGCAGGAAAAACGCGAGTGCCTGAAGACCACCTGGCAGGGT
GGGCTTCTATCCACCCATGAGGCACCTGGGCTTCTCAGCCACCCCGTGGTG
TAACAAGTGGAGGAGGAGGAGGGGGTGAGAACACCTTTCCTCCAGGCCCT
GTGCCACACTCCTTCTATGTGTGTACATGTGCCACAGGCATCACACCTTCTT
CATTAAAAACCCAGTTTCCTGTCATTTCTGCAAAAAAA

H3 sequence (unknown)

CTAAAGGTGCCTAGAGGTGCTATAGGAAACATAGATCCAGCCAGGGCTTC
CCTAAAGCAGTGCAGCACCGGCCCAGGGCATCACTAGACAGGCCCTAATT
AAGTTTTTTTTTAAAAGCTGTGTATTTATTTTAGAATCATGTTTTTCTGTAT
ATTAACTTGGGATATCGTTAATATTTAGGATATAGATTTGAGGTCAGCCAT
CTTCAAAAAAAAAAAAAAAAAAA

No sequence homology was found for the H3 cDNA in the GenBank data base. The 0.7 kb sequence of D3 includes the last two exons f *α*-

cardiac actin which include the highly conserved 200 bp UTR. The 0.3 kb F1 cDNA contains only non-coding sequences of the UTR of tropomyosin. The 1.0 kb G12 cDNA includes almost all of the coding sequences as well as the entire UTR of troponin I. The 3'UTR has been shown to function independent of the coding sequences (shown above). H3 has not been identified by sequence analysis, however, it includes a polyadenylation signal and a poly-A tail indicative of the 3' end of a cDNA. K1, the negative control cDNA that was isolated by PCR from the same cells as the complementing cDNAs, but did not activate muscle specific promoters when religated into the pCDM8 vector and introduced alone, is a full-length 0.7 kb cDNA of mitochondrial ATPase subunit 6 (Anderson et al., 1981, *supra*). Thus, the complementing cDNAs recovered from the test cells contained either all, some, or no coding sequences. The only features which they shared in common were that they were muscle structural genes and each contained sequences from the UTR.

Activity of cDNAs Maps to Sequences Within the UTR

We tested whether the UTR sequences alone, the feature common to all of the complementing cDNAs, was sufficient to induce expression of myogenic genes in NMU2. For this purpose, the troponin I and tropomyosin cDNAs were analyzed as two parts, sequences 3' and 5' to the translation termination codon, respectively. These cDNA segments were generated by PCR, ligated into the pCDM8 expression vector, and tested independently for their ability to activate the reporter gene in the NMU2-MYGpuro.1 test cell. The activity of both of these structural genes mapped to sequences within the UTR (Table 4A). Similarly, when segments of troponin I were assayed for their ability to activate the endogenous myogenin gene, the 3' segment alone exhibited activity, whereas the 5' segment did not (Table 4B). The two-fold increase in efficiency of the 3' segment alone relative to the entire G12 (troponin I) cDNA sequences may be due to enhanced expression of the shorter cDNA (0.3 versus 1.0kb). These results were confirmed by Northern analysis. In four independent clones expressing the 5' segment of tropomyosin, myogenin mRNA expression was not induced.

TABLE 4

THE ACTIVITY OF cDNAs MAPS TO THE UTRS

A. Activation f MYG-puro construct by the UTR	
Transfected cDNA*	**Puromycin^r colonies/Transfection**
Exp. 1	
Control cDNAs:	
puro	93
K1	1
Tropomyosin:	
3′	21
5′	1
Exp. 2	
Control cDNAs:	
puro	30
K1	0
Troponin:	
3′	70
5′	3

B. Activation of endogenous myogenin by the UTRs	
Transfected cDNA#	**Myogenin Positive Colonies (%)**
Control cDNAs:	
K1	1 ± 1
Troponin I:	
3′	43 ± 5
5′	1 ± 1
full length	20 ± 4

NOTES TO TABLE 4:

* Duplicate dishes containing $1x10^5$ NMU2-MYGpuro.1 c lls were transfected with 10 μg of each plasmid and selected in media containing 0.8 μg/ml puromycin (Exp. 1) or 1.0 μg/ml puromycin (Exp. 2). Shown is the total number of puromycin resistant colonies obtained after 10 days of selection. All cDNAs use the CMV promoter for expression except for the puro cDNA (puromycin N-acetyltransferase) which uses the SV2 promoter. 3' and 5' designate sequences downstream and upstream of the translation termination codon, respectively. Full length is the G12 cDNA in its entirety.

\# NMU2-HCAgpt.3 cells were co-transfected with 8 μg of cDNA plasmid and 2 μg of SV2-puro plasmid. Puromycin-resistant colonies were grown under selection with 1 μg/ml puromycin for 7 days and then maintained in medium containing 0.5% horse serum for an additional 2 days. The frequency of myogenin-positive colonies was determined following immunostaining for myogenin protein and was corrected in each experiment for backgrounds by subtracting the frequency obtained with vector alone. Approximately 100 colonies were scored for each cDNA transfected. Values are shown with standard error of the proportion.

To determine whether the activation of myogenic genes is specific to certain UTRs, we tested the activity of additional UTRs in NMU2 cells. For this purpose the UTRs of jun B, oct 1, and histone H3.3 were selected, because they encode transcripts of similar size to those of the complementing cDNAs, yet these transcripts are not differentiation-specific and are widely expressed in a number of cell types (Table 3). In order to facilitate a comparison of the activity of the different UTRs, they were all cloned into the pCDM8 vector which uses the CMV promoter to drive expr ssion. Thus, the rate of transcription of each cDNA should be

comparable. Unlike the UTRs of the complementing cDNAs, the control UTRs did not enhance myogenin expression significantly. Northern analysis r vealed that the maximum l vels of tropomyosin, troponin I and α-cardiac actin UTR RNA accumulation in stably transfected NMU2 cells were comparable to those of the UTRs expressed in primary human muscle cells. Thus, the level of UTR RNAs expressed in the stably transfected cells are achieved by human muscle cells in the course of normal myogenesis. Similar Northern analysis for control UTRs in stably transfected cells show that after a 3 day period accumulated RNA is comparable to the muscle UTRs, although of course, this quantitation is not exact due to differences among the probes. These results indicate that the observed effects are not due to toxicity resulting from over-expression of a single UTR and that the effects are specific to a subgroup of UTRs.

UTRs Increase Myogenin Expression in Wild-type Muscle Cells

To determine whether expression of the tropomyosin and troponin I UTRs could affect differentiation of muscle cells other than mutant NMU2 cells, they were transfected into the wild-type C2F3 muscle cells from which the mutant was derived. Each cDNA was transfected in duplicate in two separate experiments and at least 240 colonies of 25-200 cells were scored for myogenin expression. After 7 days of growth at clonal density in high serum medium, approximately 20% of C2F3 colonies expressed myogenin due to spontaneous differentiation. The transfection of UTR sequences derived from jun B, oct 1 or histone H3.3 did not alter the frequency of myogenin expression. By contrast, sequences derived from the UTRs of tropomyosin and of troponin I increased the frequency of positive colonies to 50-60%, two to three-fold above background. The lack of augmentation of myogenin expression by the control UTRs provides further evidence for the specificity of the effect seen with the differentiation-specific UTRs. In addition, the augmentation of myogenin expression observed in C2F3 cells upon expression of the troponin I and tropomyosin UTR sequences demonstrates that the pathway targeted by UTRs operates in a non-mutant

myoblast cell. Thus, this pathway is likely to play a regulatory role in normal myogenic differentiation.

UTRs Inhibit Growth of Fibroblasts

To determine whether the UTR sequences could induce muscle gene expression in a non-myogenic cell type, fibroblasts were transfected. 10T½ fibroblasts were selected for this purpose, because they are not myogenic but have the potential to display a myoblast phenotype upon expression of myogenic helix-loop-helix transcription factors (Davis et al. (1987) *Cell* 51, 987-1000; Edmondson and Olson (1989) *Genes Dev.* 3, 628-640). 10T½ fibroblasts were co-transfected with 2 μg of SV2-neo and 8 μg of a construct containing the Tm UTR (F1 cDNA) under control of a metallothionein promoter. Stable transfectants were selected in medium containing G418. Total RNA from the two clones, each growing in 5% dialyzed calf serum with or without 20 μM zinc, was analyzed by Northern blot hybridization with the F1 probe. Relative amounts of RNA loaded in each lane were visualized by ethidium bromide staining of the gel. 200 cells from each of two clones were seeded onto two dishes and grown in medium containing 10% calf serum for 3 days. Medium was replaced with 2% dialyzed calf serum with or without 20 μM zinc for an addition of four days of incubation, after which colonies were fixed and stained with 2% methylene blue in 50% ethanol and photographed. Although neither myogenin nor myosin heavy chain was detected, there was a significant reduction in the frequency of stably transfected 10T½ clones upon expression of the tropomyosin UTR sequences. Ten stable 10T½ transfectants were isolated by selection in G418 following co-transfection of the inducible constructs with SV2-neo. Northern analysis revealed that clone 7 did not accumulate tropomyosin UTR RNA and its growth was not affected by zinc induction. By contrast, clone 1 accumulated detectable levels of UTR RNA in the absence of zinc, and this level could be induced approximately 3-fold. At clonal density, even the basal level of expression of the UTR in clone 1 led to a detectable impairment of growth relative to the clone 7 control, but gr wth inhibition was pronounced following exposure to zinc. Although the number of colonies

appeared similar on each dish, the colony size of clone 1 was substantially smaller than the c lony size of clone 7. These experiments demonstrate that the effect of th UTR on growth corr lates with the l vel of accumulated RNA and may be dose-dependent.

To analyze the effect of the inducible constructs on 10T½ cell growth in a quantitative manner, an assay of cell number based on MTT metabolism was performed (Mosmann (1983) *J. of Immun. Methods* 65, 55-63). In this experiment, 500 cells were plated in replicates of four with or without zinc-supplemented medium for a period of four days. Cell proliferation was inhibited by 42% upon induction with zinc. These results show that expression of the UTR in fibroblasts suppresses growth.

The following study reports the results of UTRs on tumor suppression.

Suppression of Anchorage Independence

Anchorage-independent growth in the semisolid medium of soft agar is a strong indicator of the transformed phenotype in rodent cell lines. To assess the effect of Tm 3'UTR RNA expression on the transformed phenotype, stably transfected NMU2 cells were analyzed for growth in soft agar. Untransfected controls included nontransformed C2F3 cells, NMU2 and a subclone NMU2 HCA-*gpt*.3 See prior description and Rastinejad and Blau (1993) Cell 72, 903-917. Cells were seeded in parallel in liquid tissue culture medium and in medium containing soft agar. To assess viability, the number of colonies that grew in 10% calf serum in liquid culture medium was determined in each case. Under these conditions, Tm transfectants were morphologically indistinguishable from and grew as well as untransfected NMU2 cells. To assess relative anchorage independence, for each cell type the number of colonies that grew in soft agar was determined and normalized to the frequency of colony formation in liquid medium.

In three separate experiments, transformed NMU2 cells produced anchorage-independent colonies at a much greater frequency than the nontransformed C2F3 cells (Table 5). Like C2F3 myoblasts, the four Tm 3'UTR-expressing clones tested did not grow well in soft agar. By contrast, three independently derived stable clones expressing the control cDNA A6,

or clones expressing the control 3'UTR cDNAs, *junB*, and Histone H3.3, grew as well as NMU2 cells in soft agar. Northern analysis of the levels of mRNA expression for these stable transfectants were performed and the relative levels of all clones are indicated in Table 5. The data show a marked change in the anchorage-independent phenotype upon transcription of Tm but not control cDNAs expressed from the pCDM8 vector. Taken together, these results support the conclusion that suppression of anchorage independence is specific to the tropomyosin 3'UTR cDNA.

Tabl 5. Suppression of Anchorage Independence of NMU2 by Tm 3'UTR

Cells	RNA	Growth in Soft Agar Level[a]	(%)[b]
untransfected			
	C2F3		0.02
	NMU2 HCA-*gpt*.3[c]	26.3	
transfected NMU2			
	Tropomyosin 3'UTR		
	Tm clone 1	+ + +	0.4
	Tm clone 2	+ +	0.03
	Tm clone 34	+ +	1.1
	Tm clone 40	+ + +	0.01
	Control cDNAs		
	A6 clone 2	+ +	35.0
	A6 clone 3	+ + + +	20.2
	A6 clone 4	+ + +	20.6
	junB clone 5	+ + + +	14.0
	Histone clone 3	+ + + + +	21.3

[a]Relative amounts of accumulated RNA are based on Northern blot analysis.

[b]Colony formation in soft agar was corrected for the number of viable colony-forming cells in liquid medium. The results represent the average of three experiments for C2F3 and NMU2 HCA-*gpt*.3 and of two experiments for each of the transfected NMU2 clones.

[c]NMU2 HCA-*gpt*.3 is a subclone of NMU2 that was cotransfected with a silent *α*-cardiac actin-*gpt* (HCA-*gpt*) construct and SV2-*neo* and used

previously in genetic complementation experiments (Rastinejad and Blau, 1993, *supra*).

Suppression of subcutaneous tumor formation

To determine whether NMU2 cells could be used in assays of malignant transformation in vivo, cells were injected subcutaneously into athymic nude mice. The tumors that arose were analyzed histochemically and shown by a number of criteria to be rhabdomyosarcomas (Enzinger and Weiss (1988) Rhabdomyosarcome, In Soft Tissue Tumors, 2nd Ed. G. Stamithis ed., St. Louis; The C.V.Mosby Co., pp 448-488; Hiti et al. (1989) Mol. Cell. Biol. 9, 4722-4730). The malignant neoplasms contained occasional myofibers with distinct cross-striations, and an abundance of cells that expressed muscle-specific proteins such as desmin (data not shown) and sarcomeric actin. In addition, the tumors contained poorly differentiated multinuclear giant cells with darkly eosinophilic cytoplasm. In the central region of large tumors, mitotic figures were rare, whereas other regions of the tumor had a characteristically embryonal rhabdomyoblastic appearance.

To assess the effect of the 3'UTR on tumor formation in nude mice, NMU2 cells stably transfected with Tm 3'UTR, A6, or *junB* 3'UTR cDNAs were compared. The levels of RNA accumulation for these clones were determined. Untransfected NMU2 cells, the NMU2 HCA-*gpt*.3 subclone, and wild-type C2F3 cells served as controls. Various inoculum sizes of NMU2 cells were tested, and a minimum number of $1x10^5$ cells was found to be required to produce tumors consistently within 4 weeks after injection (data not shown).

The data for six separate experiments show that tumor suppression was correlated with expression of the Tm 3'UTR (Table 6). Injection of $1x10^5$ C2F3 cells did not produce tumors (0 tumors/10 sites) for up to 10 weeks, at which time the experiments were terminated. By contrast, both a stable NMU2 transfectant expressing the control cDNA A6 (clone 4) (12 tumors/12 sites) and a stable transfectant expressing the *junB* 3'UTR (12 tumors/12 sites) produced tumors in all of the injected sites within 4 weeks, a time course and efficiency consistent with that for tumor formation by

NMU2 cells (26 tumors/27 sites). Two stable transfectants that expressed relatively high levels of Tm cDNA (clones 1 and 2) showed profound suppression of tumor formation, 2 tumors/14 sites and 3 tumors/16 sites, respectively. Moreover, most tumors arising from injection of Tm-expressing clones were delayed in appearance by 2-4 weeks beyond the typical latency period seen for tumors that expressed the control cDNAs. Tm clone 3, which produced tumors in 2 of 4 sites, expressed the lowest level of Tm 3'UTR. These results indicate that suppression of tumorigenicity, like the suppression of anchorage independence, is specific to the Tm 3'UTR.

Table 6. Suppression of Tumorgenicity of NMU2 by Tm 3'UTR

Cells	RNA Level[a]	Tumor Incidence (Tumors/site)[b]
SUBCUTANEOUS		
untransfected		
C2F3		0/10
NMU2		6/6
NMU2 HCA-*gpt*.3		20/21
transfected NMU2		
Tropomyosin 3'UTR		
Tm clone 1	+ + +	2/14[c]
Tm clone 2	+ +	3/16
Tm clone 3	+	2/4[c]
Control cDNAs		
A6 clone 4	+ + + +	12/12
junB clone 5	+ + + +	12/12
INTRAMUSCULAR		
untransfected		
NMU2 HCA-*gpt*.3		13/13
NMU2		9/9
transfected NMU2		
Tm clone 1	+ + +	0/13
Tm clone 2	+ +	0/9

[a]Relative amounts of accumulated RNA are based on Northern blot analysis.

[b]Cells ($1x10^5$) were injected subcutaneously in the flanks of athymic nude mice or implanted intramuscularly in hindlimbs of *scid* and nude mice. Tumors were first detected at 3-4 weeks postinjection. Values shown

indicate tumor incidence at 10 weeks post injection for subcutaneous and 5 weeks postinjection for intramuscular sites. Data are the combined results from up to six separat experiments (subcutan ous) or two experiments (intramuscular).

[c]One tumor each that arose from Tm clone 1 and Tm clone 3 were analyzed and shown to have extinguished mRNA expression. Tumors derived from Tm clones were generally delayed in appearance.

Absence of Tm 3'UTR RNA expression correlates with loss of tumor suppression

The long latency period of tumors arising from Tm 3'UTR transfectants suggested that these tumors were derived from rare variants in which transcription of the cDNA had ceased. To test this possibility, two tumors arising following injection of Tm clone 1 and Tm clone 3 cells were examined for retention and expression of cDNA sequences. Tumors were dissected, and the dissociated cells were cultured in vitro in the presence of puromycin to restrict the analysis to the injected cells. DNA and RNA were isolated from the cultured tumor-derived cells and examined for the presence of Tm cDNA by Southern analysis and for Tm RNA expression by Northern analysis. The numbers and sizes of the hybridizing fragments in the restriction digests of the DNA analyzed by Southern blot were not altered. Thus, the cells recovered from the tumors contained Tm 3'UTR cDNA sequences that did not appear to be lost or rearranged.

By contrast, Tm RNA was undetectable in the recovered cells. These findings support the conclusion that although the transfected Tm 3'UTR cDNA was present, its expression was extinguished in these cells. For comparison, a tumor derived from cells transfected with the control cDNA A6 was analyzed. The A6 RNA accumulated as efficiently in the cells recovered from the tumor as in the cells that had been injected. From these data, it is concluded that the extinction of Tm cDNA expression in tumors is specific to that cDNA and not due to non-specific silencing of the pCDM8 CMV promoter in tumors. Moreover, th maintenance of integrated Tm cDNA in the genome of transfectants does not appear to be sufficient for tumor

suppression. Instead, it appears that the suppression of tumorigenicity is dependent on the continued expression of Tm 3'UTR RNA.

Suppression of tumor formation in muscle

To determine the fate of NMU2 cells expressing the Tm 3'UTR in the environment of the muscle, the tissue from which the cells were derived, cells were labelled with the *lacZ* gene and implanted into the hindlimb muscles of nude or *scid* mice. Within 1 week of implantation, both NMU2 cells and NMU2 cells expressing the Tm 3'UTR were detectable as β-galactosidase positive cells in a small area of the muscle. At this time, differences between legs injected with the two cell types were negligible.

Three weeks after implantation, NMU2 cells had extensively invaded the host muscle. At the site of implantation in the center of the muscle, nearly all muscle fibers had been replaced by a tumor mass. Peripheral to this site, tumor cells circumscribed virtually every muscle fiber and tissue destruction was extensive. By four weeks, tumors were detectable by palpation and after six weeks tumors were readily evident as large masses.

In contrast, at similar time points, Tm 3'UTR-expressing NMU2 cells remained highly localized at the site of implantation. Growth was minimal, and the size of the area encompassed by β-galactosidase-labeled cells after 3 weeks had not changed significantly from that observed 1 week after implantation. In addition, there was evidence of muscle regeneration with new fiber formation in the area of cell implantation. In all of the 22 sites analyzed, tumors were either undetectable or substantially delayed in appearance after injection of Tm 3'UTR expressing cells (Table 6). The finding that the NMU2 cells expressing the Tm 3'UTR survive indicates that suppression of tumor formation is due to growth inhibition, not cell death.

Neither NMU2 cells nor NMU2 cells expressing Tm 3'UTR are capable of fusion in vitro even in the presence of fusion-competent cells, but when implanted into muscle tissue both cell types were incorporated into myofibers. One week after implantation, β-galactosidase-positive myofibers were evident at implantation sites. Eventually large diameter labeled fibers of similar size to the muscle fibers of the surrounding tissue were also

apparent. The incorporation into host fibers may be due to the disruption of muscle fiber membranes at the site of implantation. These areas of muscle r generation persisted where 3'UTR expressing c lls had been injected. By contrast, 3 weeks after injection, the site of implantation of NMU2 cells was obscured by tumor growth.

IDENTIFICATION OF REGULATORY SEQUENCE OF 3' UTRS

Materials and Methods

RNAs Generated by In Vitro Transcription

The plasmids containing F1, a 214 nt segment of the *α*-Tm 3'UTR isolated by genetic complementation and controls, the ATPase A6 (700 nt; also known as K1), histone H3.3 3'UTR (500 nt), and Oct1 3'UTR (200 nt), have all been previously described (Rastinejad and Blau, 1993, *supra*). Other controls included yeast tRNA (Sigma) and poly(I):poly(C) (Pharmacia Biotechnology). For the experiments described here, these cDNAs were each subcloned from the pCDM8 vector into the pBluescript SK(-) vector (Stratagene) in the same orientation as HindIII-NotI fragments. In addition, subclones of the F1 cDNA (214 nt) were generated. Digestion of F1-pBluescript with EcoR1 and NotI, followed by a fill-in reaction with the Klenow fragment of DNA polymerase and ligation with T4 ligase produced $F1_A$, the 5' segment of F1 comprising 122 nt. Digestion of F1-pBluescript with BamH1 and HindIII, followed by fill-in and ligation, produced $F1_B$, the 3' segment of F1, comprising 111 nt. The $F1_A$ and $F1_B$ subclones contain 19 nt of overlapping sequence at their 3' and 5' ends, respectively. The orientation of $F1_A$ and $F1_B$ in pBluescript SK(-) remained the same as in F1.

For in vitro transcription reactions, F1-containing pBluescript plasmids were linearized at their 3' ends with NotI (F1 and $F1_B$) or SstI ($F1_A$). TAR-containing pBluescript plasmids were linearized with HindIII. Each linearized plasmid (1 μg) was used for in vitro transcription reactions using T7 RNA polymerase (Stratagene; Epicenter Technologies), according to the recommendations of the manufacturer. Th HIV-1 TAR sequence was derived from pLTRCAT (Kao et al., (1987) Nature *330*, 489-493) provided

by G. Nolan. The 5' 86 bp of TAR were subclon d into pBluescript SK(-) in an orientation to allow T7 RNA polymerase transcription. To generate ^{32}P-labeled transcripts, the reactions were carried out in the pr sence of 50 μCi of [^{32}P]UTP (Amersham) in a nucleotide mix containing 150 μM each of unlabeled ATP, CTP, GTP, and 15 μM UTP. For unlabeled RNA transcripts, the identical reactions were carried out in the absence of [^{32}P]UTP and in the presence of 150 μM unlabeled UTP. Such reactions yielded labeled RNA transcripts of approximately 1 x 10^8 cpm/μg. Unlabeled reaction yields were monitored by ethidium bromide staining.

Cells, RNA Dependent Protein Kinase (PKR), and Antibodies to PKR

C2F3, NMU2 and HT-1080 were cultured as described above. Cells were harvested at or near confluence and cytoplasmic extracts prepared.

NMU2 clones stably expressing $F1_A$ or $F1_B$ were derived by transfecting NMU2 HCA*gpt*.3 as previously described. Transfections were performed using the calcium phosphate method as described above. DNA precipitate containing 2 μg of selectable marker plasmid (SV2*puro*), and 8 μg of cDNA expression plasmid (in pCDM8 vector) was added to each dish. The medium was replaced 8 hr later, and following an additional 24-36 hr of incubation, cells were harvested and replated at clonal density under selection with puromycin. Clones of stable transfectants were isolated after 7-10 days using glass cloning cylinders.

Antibodies to PKR were provided by R. Petryshyn (Petryshyn et al., (1988) Proc. Natl. Acad. Sci. USA *85*, 1427-1431), I. Hovanessian (Laurent et al., (1985) Proc. Natl. Acad. Sci. USA *82*, 4341-4345), and C. Samuel (McCormack et al., (1992) Virology *188*, 47-56; Thomis et al., (1992) Virology *188*, 33-46). Rabbit polyclonal antiserum (Petryshyn et al., 1988, *supra*) was used to detect PKR from mouse-derived NMU2 cell extracts, whereas the mouse monoclonal and rabbit polyclonal antibodies that recognize human PKR were used for human-derived HT-1080 cell extracts (Laurent et al., 1985, *supra*; McCormack et al., 1992, *supra*). The IgG fraction of all of the above antisera was prepared by chromatography over protein A-sephar se (Sigma). The data obtained to establish that *α*-

tropomyosin ("*α*-Tm") induced autophosphorylation of PKR and the localization of the PKR activation to 5' fragment of *α*-Tm 3'UTR used partially purified PKR from reticulocyte lysate (Petryshyn et al.,(1983) Meth. Enzymol. *99*, 346-363) provided by R. Petryshyn and were confirmed, in part, using recombinant PKR provided by M. Katze and G. Barber.

EMSAs

EMSAs were used to examine RNA-protein interactions according to a modification of the procedure of Andino et al. (1990) Cell *63*, 369-380). Cytoplasmic extracts were prepared by harvesting confluent monolayers of NMU2 cells in phosphate-buffered saline with 0.5 mg/ml trypsin (GIBCO BRL) for 5 min, following which an equal volume of phosphate-buffered saline with 5 mg/ml trypsin inhibitor was added (Sigma). The resulting cell suspension was centrifuged at 1000 X g for 5 min and washed twice in phosphate-buffered saline containing 5 mg/ml trypsin inhibitor. The final cell pellet was resuspended in 5-pellet volumes of low salt buffer (10 mM HEPES [pH 7.9], 10 mM KC1, 1.5 mM $MgCl_2$, 0.5 mM dithiothreitol, 0.1 mM phenylmethylsulfonyl fluoride) on ice for 15 min, followed by Dounce homogenization of 10 strokes with a loose-fitting pestle. This step was followed by microcentrifugation at 4°C for 10 min. The resulting supernatant was subjected to a high salt extraction by the addition of 5x extraction buffer to a 1x final concentration (1x extraction buffer is 20 mM HEPES [pH 7.9], 550 mM NaCl, 1.5 mM $MgCl_2$, 0.5 mM dithiothreitol, 0.5 mM phenylmethylsulfonyl fluoride, 1 *μ*g/ml leupeptin, 2 *μ*g/ml antipapain, and 10 *μ*g/ml benzamidine). The high salt extract was left on ice for 10 min and then subjected to a 1 hr centrifugation at 150,000 x g. The supernatant from this spin was then desalted over P6DG resin (Bio-Rad) equilibrated in 1x binding buffer (15 mM HEPES [pH 7.9], 15 mM KC1, 10% glycerol, 5 mM $MgCl_2$, 0.5 mM dithiothreitol, 0.1 mM phenylmethylsulfonyl fluoride), and the concentration of protein in the excluded fraction was determined using Bradford reagent (Bio-Rad). These cytoplasmic extracts were then aliquoted and frozen at minus 70 C until use.

RNA-protein reactions were carried out as follows. Labeled RNA transcript (15,000 cpm) was mixed with 5 μg cytoplasmic protein in the presence of tRNA (Sigma; 20 μg of tRNA per nanogram of [^{32}P]RNA), in 1x binding buffer in a volume of 15 μl. This mixture was incubated at room temperature for 20 min, followed by the addition of 20 U RNaseT1 (Sigma) for 30 min at 37°C to digest unbound and nonspecifically bound probe, a prerequisite for obtaining specific and competable binding with F1 RNA. Unlabeled RNA competitors, when used, were added to cytoplasmic extracts at the same time as ^{32}P-labeled transcripts, with fold excesses based on mass. The IgG fraction of antisera, when used, was added to the protein extracts just prior to addition of RNA transcripts, binding buffer, and tRNA.

For EMSA analysis, 4% polyacrylamide gels containing 0.5x TBE (1x TBE equals 90 mM Tris-borate, 20 mM EDTA [pH 8.3]), 5% glycerol were prerun at 250 V (constant voltage) in 0.5x TBE running buffer for 1 hr at 4°C prior to loading RNA-protein mixtures. All electrophoresis equipment and buffers were equilibrated at 4°C prior to use. Electrophoresis was carried out at 250 V for 2.5 hr. The gels were then dried under vacuum onto Whatman 3MM paper and autoradiographed overnight using Kodak XAR-5 film and intensifying screens.

In Vitro Kinase Assays

Autophosphorylation of latent PKR was carried out in 20 μl reactions containing 30 mM Tris-HCl, (pH 7.7), 2 mM magnesium acetate, 70 mM KCl, 10 μCi of [α-^{32}P]ATP (4,500 Ci/mmol), and 5-15 ng of immunoaffinity-purified reticulocyte PKR (gift of R. Petryshyn) or 0.5 ng recombinant PKR (gift of G. Barber and M. Katze) and the indicated amounts of synthetic poly(I):poly(C) (Pharmacia) or in vitro transcribed RNAs. Reactions were incubated at 30°C for 20 min and terminated by the addition of SDS sample buffer and heating at 95°C for 5 min. Samples were then assayed by electrophoresis on 10% SDS-polyacrylamide gels, followed by autoradiography according to Petryshyn et al. (1983), *supra*.

Anchorage-Independent Assay

To assess anchorage-independent growth, cells were suspended in agar (Difco) according to the procedure described by Bouck and di Mayorca (1982) Mol. Cell. Biol. *2*, 97-105). Each cell type was tested in duplicate for growth in liquid medium and growth in soft agar. To assess viability, 200 cells were plated in duplicate dishes of liquid medium. This assessment of cloning efficiency of each cell type in liquid medium allowed correction for viability of the cells at the time of plating in all assays of cloning efficiency in agar. For determining cloning efficiency in agar, 50,000 cells were plated in 1.5 ml of Dulbecco's modified Eagle's medium containing 8% calf serum and 0.34% agar. This suspension was layered over 5 ml of a base layer of solidified Dulbecco's modified Eagle's medium containing 8% calf serum and 0.5% agar. After each week of culture, 0.5 ml of liquid medium was layered over the agar cultures to prevent desiccation. The number of large colonies (approximately >0.1 mm in diameter) with dense centers were scored for each plate after 3 weeks.

For Northern analysis of stable transfectants, total RNA was isolated by centrifugation of the cell lysates over a cesium chloride cushion (Chirgwin et al., (1979) Biochemistry *18*, 5294-5299). RNA samples were electrophoresed through 1% agarose-formaldehyde gels and transferred to Nytran hybridization membranes (Micron Separations Incorporated) using a vacuum blotter (LKB Products). The RNA was cross-linked to the membrane by exposure to ultraviolet light in a Stratalinker (Stratagene) and hybridized to the probes in a phosphate buffer as previously described (Peterson et al., (1990) Cell *62*, 493-502). $F1_A$ and $F1_B$ cDNA probes were isolated from the pCDM8 vector by XhoI digestion and separation of the insert fragment on a 2% low melting temperature agarose gel (Rastinejad and Blau, 1993, *supra*). In addition, to allow direct comparison, a probe was used that was common to all pCDM8-initiated transcripts was the 1.6 kb HincII fragment of pCDM8, which includes the transcribed vector sequences present at the 3' and 5' end of the transcripts. Approximately 25 ng of DNA was labeled with [α-^{32}P]dCTP (Amersham) using random hexamers as primer (multiprime kit, Amersham).

RESULTS

Specific Interaction of PKR with α-Tm 3'UTR RNA

As shown previously, the 3' noncoding regions of certain messenger RNAs (mRNAs) act in trans to potentiate some aspects of differentiation (Rastinejad and Blau, 1993, *supra*). The most potent of these 3'UTRs, the α-Tm 3'UTR, was characterized further. When expressed from an inducible promoter, the α-Tm 3'UTR inhibited the growth of fibroblasts. Further evidence for the involvement of the α-Tm 3'UTR RNA in growth control was derived from studies showing that its expression could suppress anchorage-independent growth and tumor formation. In all of these experiments, F1 was used, a cDNA derived from the α-Tm 3'UTR that was isolated by genetic complementation. F1 RNA (214 nt in length) initiates 367 nt downstream of the termination codon of human α-Tm. A search of the database revealed that the 5' half of F1 (designated $F1_A$) exhibited a high degree of evolutionary conservation, which was unexpected given the degree of divergence characteristic of most 3'UTRs. A prediction of RNA secondary structure was determined by free energy minimization according to the program of Zuker (Zuker, (1989) Science *244*, 48-52; Jaeger et al., (1989) Proc. Natl. Acad. Sci. USA *86*, 7706-7710) viewed with LoopViewer (Gilbert, (1990) a Macintosh program for visualizing RNA secondary structure. Published electronically on the Internet, available via anonymous ftp to iubio.bio.indiana.edu.) This particular structure forms a relatively stable stem loop with a duplex of approximately 30 bp that resembles structures generated for viral and synthetic RNAs that are activators of PKR (Mathews and Shenk, (1991) J. Virol. *65*, 5657-5662; Manche et al., (1992) Mol. Cell. Biol. *12*, 5238-5248). This structure and the similar biological functions ascribed to PKR and to the α-Tm 3'UTR supports the conclusion that the mechanism by which the α-Tm 3'UTR acts is via PKR.

To examine the interaction of the α-Tm 3'UTR RNA with PKR in cell extracts, electrophoretic mobility shift assays (EMSAs) were performed. Labeled and unlabeled F1 RNAs were derived from cDNAs subcloned into the pBluescript SK(-) vector and transcribed in vitro using T7 RNA polymerase. Cytoplasmic proteins were prepared from confluent monolayers of

nondifferentiating mutant mouse myogenic NMU2 cells, the cell type in which F1 was identified as a regulator by genetic complementation (Rastin jad and Blau, 1993, *supra*). When in vitro synthesized ^{32}P-labeled F1 RNA was incubated with cytoplasmic extracts from NMU2 cells, RNA-protein complexes were evident as two strong and several weaker bands, whereas free RNA probe migrated to the bottom of the gel. The immunoglobulin G (IgG) fraction from immune serum raised against PKR formed a ternary complex with labeled F1 RNA and cytoplasmic extracts that appeared as a supershift. By contrast, the IgG fraction from preimmune serum did not form such a complex. A 2-fold increase in the amount of either preimmune or immune IgG did not affect the quantity of ternary complex, suggesting that saturating concentrations of antibody had been achieved. Probe incubated with either preimmune or immune IgG in the absence of cytoplasmic protein extracts did not exhibit altered mobility, demonstrating that the supershifted complex required the interaction of F1, PKR, and antibody. The position of the F1-PKR complex in the absence of antibody was not apparent from experiments with crude cytoplasmic extracts, possibly because F1-PKR complexes comigrated with and were obscured by one of the more robust complexes. This finding is not unexpected given the low abundance of PKR in the cell and suggests that other as yet unidentified proteins also interact with F1 RNA.

To determine the specificity of this ternary complex, competition experiments were performed. The supershifted complex detected in the presence of antibody described above was efficiently competed by F1 RNA. Binding of PKR to RNA is not sequence specific, but is dependent on dsRNA secondary structure. Thus, a synthetic dsRNA polymer, poly(I):poly(C), which is known to bind PKR efficiently, was used as a competitor. Upon incubation with poly(I):poly(C), the supershifted RNA-protein complex was no longer apparent. As another control for specificity, TAR RNA was used, an 86 bp dsRNA segment from the 5′ end of HIV-1 RNA, known to bind PKR (Roy et al., (1991) J. Virol. *65*, 632-640). When incubated with NMU2 extracts and antibodies to PKR, a supershifted complex was obs rved with the TAR RNA that migrated at a position similar to that observed for F1 RNA.

These results suggest that PKR interacts specifically with regions of F1 dsRNA.

The results described above for antibodies to mouse PKR in mouse cell extracts were confirmed using two other antibodies that recognize human PKR in human cell extracts. Cytoplasmic extracts were obtained from HT-1080 cells, a human fibrosarcoma-derived line, incubated with the F1 probe and either monoclonal or polyclonal antibodies to human PKR and analyzed by EMSA (Laurent et al., 1985, *supra*; McCormack et al., 1992, *supra*; Thomis et al., 1992, *supra*). The complexes obtained with HT-1080 extracts exhibited similar mobilities to those obtained with NMU2 extracts, although their relative intensities differed. For each antibody, a supershifted complex was observed. These findings indicate that the F1-PKR interaction is detectable with extracts from two different cell types using three different antibodies. From these experiments, it appears that the F1-PKR interaction is not cell type or species specific. These data suggest that a specific ternary complex is formed with antibodies to PKR, F1 RNA, and PKR present in cytoplasmic cell extracts.

Activation of PKR by *α*-Tm 3'UTR RNA

To determine whether the RNA derived from the *α*-Tm 3'UTR could function as a dsRNA that activates PKR, in vitro kinase assays were performed. The binding of PKR to double-stranded regions of viral RNAs is known to induce autophosphorylation of the kinase that activates the enzyme. To determine if the *α*-Tm 3'UTR could activate PKR, RNA produced by T7 transcription from the F1 cDNA or controls was incubated at different concentrations with partially purified PKR and [α-^{32}P]ATP and analyzed by electrophoresis on SDS-polyacrylamide gels according to established procedures (Petryshyn et al., (1983) Meth. Enzymol. *99*, 346-363). Both the concentration of dsRNA required for autophosphorylation and the degree of phosphorylation are indicators of the efficacy of an RNA as an activator of PKR. Indeed, the greater the phosphate content of PKR, the more effective the enzyme is at catalyzing the phosphorylation of its substrate (Galabru and Hovanessian, (1985) Cell *43*, 685-694).

Autoradiograms of gels for each RNA analyzed covered a range of concentrations that spans nearly four orders of magnitude. The bell-shaped dose r spons is a hallmark of PKR activation that was readily apparent in these experiments. With poly(I):poly(C), a synthetic dsRNA routinely used as a positive control in PKR assays, activation of the kinase was observed at concentrations ranging from 2 to 60 ng/ml, in good agreement with the findings of others (Roy et al., 1991, *supra*). PKR activation by F1 was detectable at a similarly low RNA concentration of 0.2 ng/ml, peaked at a 10- to 100-fold higher concentration, and declined when the activating RNA was present in excess (>200 ng/ml). By contrast, control RNAs exhibited no significant activity in the kinase assay at these concentrations. These controls included tRNA that has extensive secondary structure, the ATPase A6, and the RNA derived from the 3'UTRs of histone H3.3 and Oct1 used previously as controls in the studies of growth, differentiation, and tumor suppression (Rastinejad and Blau, 1993, *supra*). Although the Oct1 and histone H3.3 3'UTRs exhibited some activity in the kinase assay, this activity was only apparent at very high RNA concentrations. These results suggest that the Tm 3'UTR RNA can activate PKR.

Computer analysis revealed that a segment of the F1 RNA of the *α*-Tm 3'UTR isolated by genetic complementation exhibited a surprisingly high degree of evolutionary conservation Table 7, given the divergence characteristic of most 3'UTRs. By RNA folding predictions, one of the structures this 122 nt RNA segment could form was a double-stranded stem loop with a duplex of approximately 30 bp (Table 8). This structure was similar to the optimal structure for dsRNA PKR activators deduced from studies of adenovirus or for HIV-1 TAR, among the best-studied viral dsRNAs that bind the kinase (Mathews and Shenk, 1991, *supra*; Roy et al., 1991, *supra*). Moreover, the properties ascribed to the *α*-Tm 3'UTR paralleled the known properties of PKR. Like the Tm 3'UTR RNA, PKR has been implicated as a regulator of cell growth with a role in preventing tumor formation (Koromilas et al., (1992) Science *257*, 1685-1689; Meurs et al., (1993) Proc. Natl. Acad. Sci. USA *90*, 232-236). In addition, PKR activity incr ases as adipocytes or myoblasts become confluent, withdraw from the cell cycl ,

and begin differentiation (Petryshyn et al., 1984, *supra,* (1988), *supra*; Birnbaum et al., 1990) Differentiation *45*, 138-145) and in adipocytes is correlated with th appearance of a poly(A)$^+$ dsRNA (Li and Petryshyn, (1991) Eur. J. Biochem. *195*, 41-48). Finally, expression of human PKR in yeast stops the cells from growing (Chong et al., (1992) EMBO J. *11*, 1553-1562). Consistent with these other discoveries concerning activation of PKR, the subject results support the role of the *α*-Tm 3'UTR acting as a cellular PKR activator.

Although several viral dsRNA activators of PKR are known, no endogenous dsRNAs with this function have previously been characterized. The above results support the use of 3'UTRs of contractile proteins for regulation of growth and differentiation.

Although eIF-2 is considered to be the preferred target of PKR in translational control, there are likely to be other targets of PKR phosphorylation. One such likely target is I*κ*B. Apparently NF-*κ*B can be activated by dsRNA in intact cells (Visvanathan and Goodbourn, (1989) EMBO J. *8*, 1129-1138). Thus, PKR likely phosphorylates members of the I*κ*B family of inhibitors that control NF-*κ*B (Ghosh and Baltimore, (1990) Nature *344*, 678-682). This could lead to an altered pattern of gene expression and loss of growth control via NF-*κ*B. In this case, the primary effect of dsRNA-induced activation of PKR on tumor suppression would be at the level of transcription, not translation. That PKR will have other targets is supported by the finding that when expression of a mutant PKR induced tumorigenicity, the levels of phosphorylation of eIF-2*α* were not decreased (Meurs et al., (1993) *supra*). Thus, other as yet unidentified targets of the serine/threonine kinase PKR may also play a role in its pleiotropic effects on growth and differentiation.

The above results show that untranslated regions, particularly of genes associated with cell division or cell differentiation, illustrated by muscle structural proteins, more particularly cytoskeletal proteins, are involved in regulation of cell division and cell differentiation. Thus, by controlling the presenc of these sequences, cells can be directed to divide or to differentiate in accordance with a predetermined schedule. By varying the

intracellular concentration of the regulatory 3'UTR, cell differentiation or growth may be selected. Employing promoters which are constitutive or inducible, th cells can be directed in a particular pathway initially, r upon induction. Furthermore, th sequences can be used in conjunction with neoplastic cells to substantially reduce the proliferative capability of the neoplastic cells. By employing the UTR sequences of the subject invention, either individually or in combination, or employing sequences in the antisense direction or as part of a ribozyme, cell regulation can be achieved in a variety of environments, both *in vitro* and *in vivo*.

All publications and patent applications cited in this specification are herein incorporated by reference as if each individual publication or patent application were specifically and individually indicated to be incorporated by reference.

Although the foregoing invention has been described in some detail by way of illustration and example for purposes of clarity of understanding, it will be readily apparent to those of ordinary skill in the art in light of the teachings of this invention that certain changes and modifications may be made thereto without departing from the spirit or scope of the appended claims.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: Blau, Helen
Rastinejad, Farzan

(ii) TITLE OF INVENTION: Cellular Regulation with Riboregulators

(iii) NUMBER OF SEQUENCES: 18

(iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: Flehr Hohbach Test Albritton & Herbert
(B) STREET: Four Embarcadero Center, Suite 3400
(C) CITY: San Francisco
(D) STATE: CA
(E) COUNTRY: USA
(F) ZIP: 94111

(v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.25

(vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER: PCT/US94/
(B) FILING DATE: 23 March 1994
(C) CLASSIFICATION:

(viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: Rowland, Bertram I
(B) REGISTRATION NUMBER: 20015
(C) REFERENCE/DOCKET NUMBER: FP57973-1/STAN-149-1

(ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: 415-781-1989
(B) TELEFAX: 415-398-3249

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 36 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

GAGACCGGAA GCTTCTAGAG ATCCCTCGAC CTCGAG 36

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 36 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

TCTAGAGTCG CGGCCGCGAC CTGCAGGCGC AGAACT 36

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

TCGACCTCGA GATCCATTCT 20

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

TAGGTATGGA AGATCCCTCG 20

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

TCACATGTTG TTTAACTCCA GTAA 24

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

CCCTCGAGCT GGGCAGAGCT GG 22

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 31 base pairs
(B) TYPE: nucleic acid

(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

CCCCCCGCGG CCGCTTTATT ACAAAATTAG G 31

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

GGCTCGAGAA CGTCCCCTGC CCCT 24

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 35 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

CCCCGCGGCC GCTTAAATAG ATTCAATAAA AAGAA 35

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

TTGCGGCCGC TAGGTATGGA AGATCCCTCG 30

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 199 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

CGCAAATGCT TCTAAGATGC CTTCTCTCTC CATCTACCTT CCAGTCAGGA
TGACGGTATT 60

ATGCTTCTTG GAGTCTTCCA AACCACCTTC CCTCATCTTT CATCAATCAT
TGTACAGTTT 120

GTTTACACAC GTGCAATTTG TTTGTGCTTC TAATATTTAT TGCTTTCATA
AATAAACCAG 180

ACCAGGACTT GCAACCTAA 199

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 214 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

CAAAAGGTCC CCCTGTGGTC TTTTGTGTCA ACATTGTACA ATGTAGAACT
CTGTCCAACA 60

CTAATTTATT TTGTCTTGAG TTTTACTACA AGATGAGACT ATGGATCCCG
CATGCCTGAA 120

TTCACTAAAG CCAAGGGTCT GTAAGCCACG CTGCTCTTCT GAGACTTCCA
ATCAGAAAGG 180

AATGGAAGTC TCAGAGAGCC AGCGTGGCTT ACAG 214

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 292 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

CTTCTTCCCT CCAGCCTGCA ATGCCCTCCT CTGGAACTGG GATTAAACAG
ATACCCAAGA 60

GGCAGGAAAA ACGCGAGTGC CTGAAGACCA CCTGGCAGGG TGGGCTTCTA
TCCACCCATG 120

AGGCACCTGG GCTTCTCAGC CACCCCGTGG TGTAACAAGT GGAGGAGGAG
GAGGGGGTGA 180

GAACACCTTT CCTCCAGGCC CTGTGCCACA CTCCTTCTAT GTGTGTACAT
GTGCCACAGG 240

CATCACACCT TCTTCATTAA AAACCCAGTT TCCTGTCATT TCTGCAAAAA AA 292

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 226 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

CTAAAGGTGC CTAGAGGTGC TATAGGAAAC ATAGATCCAG CCAGGGCTTC CCTAAAGCAG 60

TGCAGCACCG GCCCAGGGCA TCACTAGACA GGCCCTAATT AAGTTTTTTT TTAAAAGCTG 120

TGTATTTATT TTAGAATCAT GTTTTTCTGT ATATTAACTT GGGATATCGT TAATATTTAG 180

GATATAGATT TGAGGTCAGC CATCTTCAAA AAAAAAAAAA AAAAAA 226

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 120 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

CAAAAGAACC CCTGGTGTCT TTTGTTCAGT CTAATACAAT TTAGACTTCT GTCCAACACT 60

AATTTATTTT TGTCTTGAGT AGCACTGCGA GAGGATATGG GTTCCAATGT GCCCCGTGTT 120

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 120 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

CAAAAGGGCC TCTGGGGTCT TTTGTTCAAT ATTTTACAAT TTAGACTTCT GTCCAACACT 60

AATTTATTTT GTCTTGAGTG TACTGCGAGA GAAGAATATG GGGTTTTGTA TGCCCGTGTT 120

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 115 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

CAAAGGTCCC CCTGTGGTCT TTTGTTCAAC ATTGTACAAT GTAGAACTCT GTCCAACACT 60

AATTTATTTT GTCTTGAGTT TTACTACAAG ATGAGACTAT GGATCCCGCA
TGCCT 115

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 121 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

CAAAAGGTCC CCCTGTGGTC TTTTGTGTCA ACATTGTACA ATGTAGAACT
CTGTCCAACA 60

CTAATTTATT TTGTCTTGAG TTTTACTACA AGATGAGACT ATGGATCCCG
CATGCCTGAA 120

T 121

WHAT IS CLAIMED IS:

1. A cell, or progeny thereof, into which has been introduced a cell growth modulating amount of: (1) an untranslated region ("UTR") of a cellular genomic sequence under the transcriptional regulation of a promoter; (2) the complementary sequence of said UTR; or (3) a ribozyme comprising an homologous sequence of said UTR;

said genomic sequence being transcribed during cell division or differentiation, wherein said UTR is substantially free of coding sequences and modulates at least one of cell division or cell differentiation.

2. A cell according to Claim 1, wherein said UTR specifically binds to double-stranded RNA-dependent protein kinase.

3. A cell according to Claim 1, wherein said UTR is a 3'UTR.

4. A cell according to Claim 3, wherein said UTR is from a gene encoding an adhesion protein, a structural protein, a contractile protein, an enzyme, a regulatory protein, a growth factor protein, a receptor, or an oncogene.

5. A cell according to Claim 4, wherein said genomic sequence encodes a cytoskeletal protein.

6. A cell according to Claim 5, wherein said UTR is from a tropomyosin, troponin I or α-cardiac actin.

7. A cell according to Claim 1, wherein said UTR is dsRNA of at least about 60nt and is a fragment of a naturally occurring 3'UTR which forms a stem and loop secondary structure.

8. A cell according to Claim 1, wherein said UTR or complementary sequence is transcribed as a result of a construct integrated into a chromosome of said cell.

9. A cell according to Claim 1, wherein said cell is a neoplastic cell.

10. A cell according to Claim 9, wherein said UTR or complementary sequence is from an extracellular source.

11. A cell according to Claim 9, wherein said UTR or complementary sequence comprises a regulatory portion of the *α*-tropomyosin 3'UTR.

12. A DNA sequence comprising a transcriptional initiation region functional in a mammalian cell and a sequence encoding a UTR substantially free of coding sequence or the complementary sequence of said UTR or a ribozyme comprising a complementary sequence of said UTR, said sequence capable of specifically binding to double-stranded RNA-dependent protein kinase,

wherein said promoter is transciptionally active during cell division and/or cell differentiation in a mammalian cell.

13. A DNA sequence according to Claim 12, wherein said UTR is at least a regulatory fragment of a 3'UTR.

14. A DNA sequence according to Claim 13, wherein said UTR sequence is from an adhesion protein, a structural protein, a contractile protein, an enzyme, a regulatory protein, a growth factor protein, a receptor, or an oncogene.

15. A DNA sequence according to Claim 14, wherein said sequence is from a cytoskeletal protein.

16. A DNA sequence according to Claim 15, wherein said cytoskeletal protein is tropomyosin, troponin I or *α*-cardiac actin.

17. An isolated DNA sequence encoding at least a fragment of a 3' UTR of a cytoskeletal gene substantially free of native coding sequence, or the H3 sequence, or regulatory active fragment thereof; the complementary sequence thereof; or a ribozyme comprising a sequence homologous to at least a portion of said UTR, and capable of specifically binding to double-stranded RNA-dependent protein kinase.

18. A DNA sequence according to Claim 17, wherein said cytoskeletal gene is tropomyosin, troponin I or *α*-cardiac actin.

19. A synthetic DNA sequence comprising at least one unnatural nucleotide and comprising at least 60nt of the untranslated region ("UTR") of a genomic sequence which is naturally transcribed to RNA, or the complementary sequence of said UTR, or a ribozyme comprising an homologous sequence of said UTR, said genomic sequence being transcribed during cell division or differentiation, wherein said UTR is substantially free of coding sequences, and wherein said UTR is capable of modulating cell division or cell differentiation.

20. A synthetic DNA sequence according to Claim 19, wherein said genomic sequence encodes an adhesion protein, a structural protein, a contractile protein, an enzyme, a regulatory protein, a growth factor protein, a receptor, or an oncogene.

21. A mutated cell useful for genetic complementation of agents which regulate cell division and/or growth having the following characteristics:

mammalian;

fails to express at least one gene characteristic of differentiation;

is capable of anchorage independent growth in soft agar;

is capable of producing tumors in immunodeficient mice when introduced at not more than about 10^7 cells;

has a low reversion to the parental cell from which it was derived spontaneously or by mutagenesis; and

shows increased cell growth when fused to said parental cell.

22. A mutated cell according to Claim 21, wherein said cell is myogenic.

23. A mutated cell according to Claim 22, wherein said cell fails to express at least one gene of the helix-loop-helix family or differentiation specific structural protein family.

24. A method for detecting genetic complementation of nucleic acid agents capable of regulation of cell growth and/or differentiation, said method comprising:

introducing into a parent cell according to Claim 21 a construct comprising a transcriptional initiation region functional in said cell and a test sequence to provide genetically modified cells; and

screening said genetically modified cells comprising said construct by at least one of the following assays;

a) growing said cells in soft agar to establish anchorage independent growth; or

b) introducing a sufficient number of said genetically modified cells into an immunocompromised host, where the number of parent cells would result in a tumor, and detecting the incidence of tumors resulting from said genetically modified cells as compared to the parent cells.

25. A method for determining UTR regulatory sequences for regulating double-stranded RNA-dependent protein kinase, said method comprising:

analyzing genomic sequences encoding UTR RNA of at least about 60nt to provide a candidate RNA sequence;

combining RNA having said candidate sequence with double-stranded RNA-dependent protein kinase; and

determining the binding affinity of said RNA having said candidate sequence with said double-stranded RNA-dependent protein kinase, whereby said binding affinity is a measure of the regulatory activity of said RNA having said candidate sequence.

26. A purified chemical compound characterized by being:

capable of regulating mammalian cell differentiation;

capable of competing at the same concentration with a UTR RNA for binding to receptor to the minimal binding sequence for said UTR RNA;

capable of diminishing cell division of a neoplastic mammalian cell; and

has a molecular weight of less than about 5kD.

27. A chemical compound comprising a nucleic acid sequence having the sequence of at least 18nt of a UTR of a gene transcribed during cell differentiation and other than a housekeeping gene, or the complementary sequence thereof, covalently joined to a label capable of providing a detectable signal or binding to a receptor.

28. A chemical compound comprising a nucleic acid sequence comprising at least one unnatural nucleotide and encoding the UTR of a gene transcribed during cell differentiation and other than a housekeeping gene, said gene substantially free of native coding sequence or the complementary sequence thereof, covalently joined to a label capable of providing a detectable signal or binding to a receptor.

29. A method of inhibiting the growth of a neoplastic cell, said method comprising:

contacting said neoplastic cell with a growth inhibiting amount of a nucleic acid sequence having the sequence of a UTR of a gene transcribed during cell differentiation and other than a housekeeping gene binding specifically to double-stranded RNA-dependent protein kinase.

30. A method of inhibiting viral proliferation in or viral transformation of a cell, said method comprising:

contacting said cell with a proliferation inhibiting amount of a nucleic acid sequence having the sequence of a UTR of a gene transcribed during cell differentiation and other than a housekeeping gene binding specifically to double-stranded RNA-dependent protein kinase.

A. **CLASSIFICATION OF SUBJECT MATTER**

IPC(5) :C07H 21/00; C12N 15/00, 5/00

US CL : 536/24.1; 435/172.3, 240.2

According to International Patent Classification (IPC) or to both national classification and IPC

B. **FIELDS SEARCHED**

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 536/24.1; 435/172.3, 240.2

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

NONE

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

APS, DIALOG, MEDLINE

C. **DOCUMENTS CONSIDERED TO BE RELEVANT**

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y,P	Cell, Vol. 72, issued 26 March 1993, Rastinejad et al. "Genetic complementation reveals a novel regulatory role for 3' untranslated regions in growth and differentiation", pages 903-917, see entire document.	1-30

☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be part of particular relevance

"E" earlier document published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
26 APRIL 1994	MAY 16 1994
Name and mailing address of the ISA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231 Facsimile No. 703-305-3230	Authorized officer Gary L. Brown Telephone No. (703) 308-0196

Form PCT/ISA/210 (second sheet)(July 1992)*